Università di Pisa

Facoltà di Scienze Matematiche Fisiche e Naturali

Corso di Laurea Specialistica in Scienze Fisiche

Anno Accademico 2009/2010

Tesi di Laurea Specialistica

Polytopes and Loop Quantum Gravity

Candidato
Pietro Doná

Relatore
Prof. Simone Speziale

Relatore
Prof. Pietro Menotti

Contents

1 Introduction

General relativity and quantum theory are often claimed as the two main chapters of modern physics. General relativity and quantum theory describes perfectly all the natural phenomena we can observe. Quantum theory, in the form of strong, weak interaction and ordinary quantum mechanics and QFT, governs the range of observations from $10^{-19}m$ up to a fraction of millimeter. General relativity provides an accurate description of large scale phenomena from a millimeter to cosmological distances.

Analyzing the whole panorama from a wider prospective, it is far from obvious that the two chapters are part of the same story: the requirement of a dynamical space-time with no preferred frame, the main pillar of general relativity, clearly clashes with the quantum theory requisite of a fixed background structure and a preferred time coordinate.

Despite more than a half century of research we don't have a complete quantum theory of gravity. The failure to quantize gravity rests only in part on technical difficulties: general relativity is a complicated and highly nonlinear theory and only in 1986 it was finally shown that conventional QFT techniques fail [1]. But the real problems are almost certainly deeper: quantum gravity requires a quantization of spacetime itself, and at a fundamental level we do not know what that means.

When this thesis is completed still lacks a compelling argument that invalidates or supports any of the present theories of quantum gravity. Even an observational evidence that point us in a particular direction is still missing. Quantum gravity remains a theorists' playground, an arena for "theoretical experiments," some of them quite adventurous, which may or may not stand the test of time.

1.1 Motivation for quantum gravity

The first natural objection one can rise against quantum gravity is "Why we need to quantize gravity?". The problems addressed by general relativity and those addressed by quantum theory typically arise at very different length and energy scales, and there is not yet any direct experimental evidence that gravity is quantized. What tells us that we have to quantize it?

The first argument is to mention the unity of physics. In fundamental physics unity has been certainly an important guiding light that leads us i.e. to Maxwell's unification of electromagnetism or to the Weinberg-Salam electroweak model. But since such a historical argument is not entirely convincing we must go further.

If we suppose gravitational field is not quantized we obtain a theory that could lead to a violation of the uncertainty principle (studied by Eppley and Hannah (1977) [2] Squires and Pearle (1996)[3]), using gravitational waves one could simultaneously determine the position and momentum of a particle to arbitrary accuracy[1].

There is a second line of argument for quantizing gravity, arising more from hope than necessity. Both classical general relativity and quantum field theory have serious limitations, and there is some reason to believe that quantum gravity may offer a cure.

While no one has proven that the quantization of gravity will eliminate singularities, this is the sort of thing one might expect from a quantum theory. In particular we hope that quantum theory will remove singularities of black holes and in Cosmology. A proper treatment of quantum gravity might even determine initial conditions for the Universe, making cosmology a completely predictive science [4].

Also QFT has its own problems, in the form of the infinities that plague perturbation theory. From the modern point of view, most quantum field theories are really "effective field theories," in which the divergences reflect our ignorance of physics at very high energies. It has long been speculated that the

[1] It has, however, been criticized that the gedanken experiments presented by Eppley and Hannah (1977) cannot even be performed in principle, and even if they could, it does not necessarily follow that gravity must be quantized (Mattingly 2006).

missing ingredient is quantum gravity, which has a natural length scale and might provide an automatic cutoff at the Planck energy.

We thus approach quantum gravity with a mixture of hope and fear: hope that it can solve some fundamental problems in general relativity and quantum field theory and perhaps offer us a unified picture of physics, and fear that a failure will demonstrate underlying flaws in the physics we think we now understand.

1.2 Different approaches to Quantum gravity

There are clearly two most popular approaches to quantum gravity: a major one, string theory, popular among particle physicists, and a (distant) second, loop quantum gravity, popular among relativists. String theory [5] can be seen as the natural outcome of the line of research that started with the effort to go beyond the standard model, and went through grand unified theories, supersymmetry and supergravity. Loop quantum gravity [22] can be seen as the natural outcome of the line of research that started with Dirac's interest in quantizing gravity, which led him to the development of the theory of the quantization of constrained systems; and continued with the construction of canonical general relativity

String theory and loop quantum gravity are characterized by surprising similarities (both are based on one-dimensional objects), but also by a surprising divergence in philosophy and results.

String theory defines a superb "low" energy theory, but finds difficulties in describing Planck scale quantum spacetime directly. From the point of view of quantum gravity the main physical results of string theory are two: the derivation of the Bekenstein-Hawking formula for the entropy of a black hole as a function of the horizon area [7] and the indications that the spacetime continuum is meaningless below the Planck length. In fact in order to probe smaller distances one needs higher energy, but at high energy the string "opens up from being a particle to being a true string" which is spread over spacetime, and there is no way of focusing a string's collision within a small spacetime region. A more fashion development of string theory is the so called AdS/CFT correspondence [6] that relates a quantum field theory in one dimension to a theory in one higher dimension that includes gravity.

On the other hand the main difficulty with string theory is the lack of a complete nonperturbative and *background independent* formulation of the theory.

Loop quantum gravity provides a beautiful and compelling account of Planck scale quantum spacetime, but finds difficulties in connecting to low energy physics. As string theory also loop quantum gravity provides a computation of black hole entropy [8, 9, 10]. Moreover LQG provides also a notion of quantum geometry: the leading result is discreteness of the spectra of geometric operators (area and volume) [11] that leads to a granular interpretation of space; recently it has been proposed an interesting characterization of these grains of space as superposition of flat polyhedra[51]. Another recent important result is the asymptotic analysis of the EPRLS spin foam model (the most accredited model for LQG dynamic) for some classes of geometrical boundary data: asymptotically this model seems to reproduce Regge gravity. Finally LQG provides control of spacetime singularities, such as those in the interior of black holes and the cosmological one. The application of loop quantum gravity to cosmology is one of its most spectacular achievements.

Although LQG has a lot of interesting results, there are a lot of open problems and difficulties to solve in order to obtain a complete physical theory. The main is how to define properly the semiclassical limit in order to produce some prediction. Even if we begin to figure out what states correspond to piecewise flat discontinuous geometries, how can a smooth geometry emerge from the full theory is still not clear. Solving this problem has many facets: fully understand the dynamics of the theory, find how to couple the theory to matter fields, realize what are the right variables needed to describe a perturbative expansion. Although we have made many steps forward, the path is still long.

Besides strings and loops, a number of other approaches are being investigated. A substantial amount of energy has been recently devoted to the attempt of defining quantum gravity from a discretization of general relativity, on the model of lattice QCD (Dynamical triangulations, quantum Regge calculus, simplicial models).

We need to mention also a particular approach (called "asymptotic safety") that uses renormalization group techniques to define a non-gaussian fixed point.

A number of approaches (Euclidean quantum gravity, old perturbative quantum gravity, quantum field theory on curved spacetime) aim at describing certain regimes of the quantum behavior of the gravitational field in approximate form, without the ambition of providing the fundamental theory, even if they previously had greater ambitions.

1.3 A snapshot of Loop Quantum Gravity

Before entering in the details we give a brief overview of the state of the theory up to date.

Loop quantum gravity is, in its original version, a canonical approach to quantum gravity. Nowadays, a covariant formulation of the theory exists in the so called spin foam models. Loop quantum gravity is background independent: his construction does not depend on fixed classical geometric structures and this is the main distinct feature of the theory.

New techniques had to be developed for this, and the resulting Hilbert spaces look very different than those in standard quantum field theory, with one dimensional excitations of the fields.

Background independence makes the contact with low energy physics a complicated endeavor. The latter problem has attracted a considerable amount of work, but is still not completely solved. Another (related) challenge is to fully understand the implementation of the dynamics. In loop quantum gravity the question of finding quantum states that satisfy "quantum Einstein equations" is reformulated as finding states that are annihilated by the quantum Hamilton constraint. The choices that go into the definition of this constraint are poorly understood in physical terms. Moreover the constraint should be implemented in an anomaly-free way, but what this entails in practice, and whether existing proposals fulfill this requirement are still under debate. This is partially due to the lack of physical observables with manageable quantum counterpart, to test the physical implication of the theory.

While these challenges remain, remarkable progress has happened over the last couple of years: The master constraint program has brought new ideas to bear on the implementation of the dynamics. Progress has been made in identifying observables for general relativity that can be used in the canonical quantization. A revision of the vertex amplitudes used in spin foam models has brought them in much more direct contact to loop quantum gravity.

Black holes are fascinating objects predicted by general relativity and, because of the singularities within, they even point beyond the classical theory. Therefore they are a tempting subject of investigation in any theory of quantum gravity. Loop quantum gravity was able to successfully describe black hole horizons in the quantum theory: it is possible to identify degrees of freedom that carry the black hole entropy, and prove, for a large class of black holes, the Bekenstein-Hawking area law.

And, last but not least, in loop quantum cosmology, the application of the quantization strategy of loop quantum gravity to mini-superspace models has become a beautiful and productive laboratory for the ideas of the full theory, in which the quantization program of loop quantum gravity can be tested, and, in many cases, brought to completion.

1.4 An overview of our results

One of the major open problem in LQG is understanding how to define the semiclassical limit. One useful idea is to consider in the full theory states with support on a single graph. This truncation captures only a finite number of degrees of freedom of the theory and corresponds to approximate a smooth geometry with a discrete one.

If one wants to work with this approximation it's really useful to have a geometric intuition of the classical degrees of freedom contained. The Hilbert space of a single graph decomposes into a direct sum of intertwiner spaces, one for each node. It is already known that a four-valent node can be interpreted as a quantum tetrahedron. More precisely, the quantization of the space of shapes of a flat tetrahedron in $\mathbb{R}^3$ with fixed areas is the space of four valent intertwiner.

Why we have to study polytopes (namely convex bounded polyhedra)? What's their connection with loop quantum gravity? If we want to extend the geometric intuition of quantum tetrahedron to the generic intertwiner space, we have to find a geometrical object uniquely identified by an intertwiner. This is exactly a polytope.

The main goal of this thesis is to investigate this correspondence. In a first part we focus on characterizing classical polytopes in $\mathbb{R}^3$. We provide the necessary tools to speak about polytopes in the language of LQG: we study how to reconstruct the solid figure and how to compute geometrical quantities (the volume, the edge lengths) in terms of LQG variables. Using coherent states we explore the large spin limit of the intertwiner space. Coherent intertwiners are basically labeled by a classical polytope but this is not only an abstract correspondence, in fact they are also peaked on these discrete geometries.

In the last part of the thesis we discuss some applications. The first one is a technique that allows us to promote a classical observables on the space of polytopes to a quantum operator, with the desired semiclassical behavior, in a natural way. In particular we proposed a volume operator such as, in the large spin limit, it reproduces the geometrical volume of a polytope. This operator is bounded and positive semi-definite, but its main property is the cylindrically consistence: the operator defined on a intertwiner space with a null spin is equivalent to the operator defined on the one less valent intertwiner space without that spin. We also computed numerically its spectrum in some simple cases and we found that a gap is present.

Finally we present some implications in the study of semiclassical limit of the dynamic of LQG. An extension of the EPR spin foam model to generic valence has been proposed. No derivation of this model are given. A lesson from the recent asymptotics studies of the EPR models is that the amplitude is dominated by saddle points. When the boundary data are compatible with the existence of a (unique) flat 4-simplex, then the amplitude is effectively given in terms of the Regge action for the 4-simplex. We propose to study the large spin limit of this new model and to see if we obtain that under some condition the amplitude is related to some "effective" Regge action where the cells are our polytopes.

2 A path to loop quantum gravity

In order to motivate the loop approach, it is useful to begin with a review of the two main historical paths to the quantization of general relativity: the covariant, or functional integral, approach, and the canonical, or Hamiltonian, approach. Their partial successes and difficulties will shed light on the problem of quantum gravity.

2.1 Covariant approach

The basic goal of the covariant approach is to define a functional integral for general relativity,

$$\int \mathcal{D}g_{\mu\nu} e^{-iS_{EH}(g)}, \tag{1}$$

where

$$S_{EH}(g_{\mu\nu}) = \frac{1}{16\pi G}\int \mathrm{d}^4x\,\sqrt{-g}\,g^{\mu\nu}R_{\mu\nu}\left(\Gamma\left(g\right)\right) \tag{2}$$

is the Einstein-Hilbert action. Here g is the determinant of the metric, and $\Gamma(g)$ is the Levi-Civita connection entering the covariant derivative of vectors and tensors,

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\rho\mu}(g)v^\rho, \qquad \Gamma^\nu_{\rho\mu}(g) = \frac{1}{2}g^{\nu\lambda}[\partial_\rho g_{\lambda\mu} + \partial_\mu g_{\lambda\rho} - \partial_\lambda g_{\rho\mu}]. \tag{3}$$

From courses in QFT, we know how to properly define only the Gaussian integral

$$\int \mathcal{D}\varphi \exp\left(-\int \frac{1}{2}\varphi\Box\varphi\right). \tag{4}$$

This corresponds to a free theory. An interacting theory can be treated in perturbation theory from a generating functional based on (4). However, there is no quadratic term in the Einstein-Hilbert action (in fact, it is not even polynomial). This obstruction can be bypassed if one performs a perturbative expansion of the metric. Consider the field redefinition

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{5}$$

where $\eta_{\mu\nu}$ is fixed to be the Minkowski metric, and $h_{\mu\nu}$ the new dynamical field. The introduction by hand of the background field $\eta_{\mu\nu}$ is crucial: if we treat $h_{\mu\nu}$ as a small fluctuation, a perturbative expansion of the action gives (in the De Donder gauge)

$$S_{EH}(g_{\mu\nu}) = \frac{1}{32\pi G}\int \mathrm{d}^4x\, h_{\mu\nu}\Box h^{\mu\nu} + o(h^3). \tag{6}$$

The background metric can be also used to define the Wick rotation. This action is amenable to the machinery of QFT. Among the main results of this background-dependent perturbative approach, let us highlight the following ones:

1. The quanta of the gravitational field are interpreted as massless spin-2 particles (the "gravitons").
2. In the low energy static limit, one can compute quantum corrections to the classical Newton potential [13], recovering the classical relativistic correction $o(G^2)$ and a new, purely quantum one $o(\hbar G^2)$:

$$V_N = -G\frac{m_1 m_2}{r}\left(1 + G\frac{m_1+m_2}{r} - \frac{127}{30\pi^2}\frac{\hbar G}{r^2} + o(r^{-3})\right) \tag{7}$$

3. As for other quantum field theories, one can define the S-matrix to describe the fundamental observables, and compute scattering amplitudes of gravitons with matter fields and with themselves [14]

This brief list does not exhaust the interesting results obtained with this approach. On the other hand, there is a fundamental difficulty with it. As for any quantum field theory, infinities appear when considering the effects of arbitrarily small ("ultraviolet") field fluctuations. These divergences are usually dealt with via the procedure of renormalization. This technique, successful for $\lambda\phi^4$ or gauge theories, fails in the case of gravity. This can be expected from dimensional arguments, and the rigorous proof that the perturbative quantization of general relativity fails because of non-renormalizable ultraviolet divergences was obtained in the late eighties, by Goroff and Sagnotti [1].

This means that the theory constructed as sketched above can be used for low-energy calculations, but it would be inconsistent if taken seriously at all energy scales. Heuristically, the inconsistence can be explained as follows. The key field redefinition (5) assumes that we can quantize the field $h_{\mu\nu}$, and that its dynamics takes place on a fixed classical background $\eta_{\mu\nu}$. However, as we increase the energy of the fluctuations its backreaction increases, and it becomes inconsistent to assume that the background stays fixed, unperturbed. Although gravitons capture correctly the low-energy physics of the gravitational field, they might not be the right quantities to describe the quanta of gravity at Planckian energies.

2.2 Canonical approach

The canonical formalism is based on second quantization, and describes the quantum theory in terms of functionals of the fields, e.g. $\Psi[\phi]$ for a scalar field ϕ, like the familiar function of the configuration variables $\psi[x]$ in quantum mechanics. The dynamics is described by the quantum Hamiltonian $\hat{H}$ and the Schödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi[\phi] = \hat{H}\left(\phi, \frac{\delta}{\delta\phi}\right)\Psi[\phi] \tag{8}$$

For the reader less familiar with this approach, see [15]. Given the overwhelming success of the functional integral approach to the quantization of (non-gravitational) field theories, the canonical formalism is often underappreciated in many courses, and a certain taste of "old fashion" is associated with it. It was applied to general relativity by Arnowitt, Deser and Misner (ADM), Dirac, Wheeler and De Witt, among many others. The loop approach is related to this original idea, which we will now review in some details. For more on it, see e.g. [16].

2.2.1 ADM formalism

In order to put the Einstein-Hilbert action into canonical form one needs to identify the variables which are canonically conjugated, and then perform the Legendre transform. To that end, we make the assumption that $\mathcal{M}$ has the topology $\mathcal{M} \cong \mathbb{R}\times\Sigma$ where Σ is a fixed three-dimensional manifold of arbitrary topology and spacelike signature. This poses no restrictions if we require that $\mathcal{M}$ has no causally disconnected region (more precisely that $\mathcal{M}$ is globally hyperbolic). In fact by a theorem due to Geroch and improved by Bernal and Sanchez [17, 18] if the space-time is globally hyperbolic then it is necessarily of this kind of topology.

Having made this assumption, one knows that $\mathcal{M}$ foliates into a one-parameter family of hypersurfaces $\Sigma_t = X_t(\Sigma)$ embeddings of Σ in $\mathcal{M}$. The foliation allows us to identify the coordinate $t \in \mathbb{R}$ as a time parameter. Notice however that this "time" should not be regarded as an absolute quantity, because of the diffeomorphism invariance of the action. A diffeomorphism $\phi \in \mathrm{Diff}(\mathcal{M})$ maps a foliation X into a new one $X' = X \circ \phi$, with a new time parameter t'. Conversely, we can write a general diffeomorphism

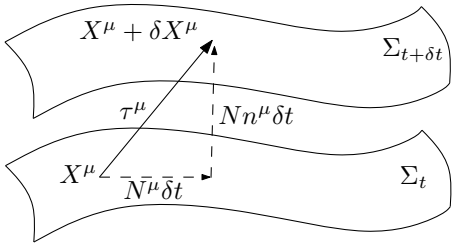
$\phi \in \mathrm{Diff}\,(\mathcal{M})$ as composition of different foliations, $\phi = X' \circ X^{-1}$. Hence, we can always work with a chosen foliation, but the diffeomorphism invariance of the theory will guarantee that physical quantities are independent of this choice. A theory of spacetime which unlike general relativity has a preferred foliation and thus a preferred time, is a theory which breaks diffeomorphism invariance.

Given a foliation X_t and adapted ADM coordinates (t, x), we can define the time flow vector

$$\tau^\mu(x) \equiv \frac{\partial X_t^\mu(x)}{\partial t} = (1, 0, 0, 0). \tag{9}$$

This vector should not be confused with the *unit normal* vector to Σ, which we denote n^μ. They are both timelike, $g_{\mu\nu}\tau^\mu\tau^\nu = g_{00}$ and $g_{\mu\nu}n^\mu n^\nu = -1$, but they are not parallel in general. Let us decompose τ^μ into its normal and tangential parts,

$$\tau^\mu(x) = N(x)n^\mu(x) + N^\mu(x). \tag{10}$$



It is convenient to parametrize $n^\mu = (1/N, -N^a/N)$, so that $N^\mu = (0, N^a)$. N is called lapse function, and N^a shift vector. In terms of lapse and shift, we have

$$\begin{aligned} &g_{\mu\nu}\tau^\mu\tau^\nu = g_{00} = -N^2 + g_{ab}N^aN^b, \\ &g_{\mu\nu}\tau^\mu N^\nu = g_{0b}N^b = g_{\mu\nu}(Nn^\mu + N^\mu) = g_{ab}N^aN^b \;\longrightarrow\; g_{0a} = g_{ab}N^b \equiv N_a. \end{aligned}$$

Using these results, the metric tensor can be written as

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = -\left(N^2 - N_aN^a\right)\mathrm{d}t^2 + 2N_a\mathrm{d}t\mathrm{d}x^a + g_{ab}\mathrm{d}x^a\mathrm{d}x^b, \tag{11}$$

where $a = 1, 2, 3$ are spatial indices and are contracted with the 3-dimensional metric g_{ab}.

Notice that the spatial part g_{ab} is *not* in general the intrinsic metric on Σ_t. The latter is given by

$$q_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu. \tag{12}$$

Tensors on the spatial slide Σ, since their scalar product with n vanishes, can be equivalently contracted with g or q. The quantity $q^\mu_\nu = g^{\mu\rho}q_{\rho\nu}$ acts as a projector on Σ_t, allowing us to define the tensorial calculus on Σ_t from the one on $\mathcal{M}$. An important quantity is the extrinsic curvature of Σ_t,

$$K_{\mu\nu} = q^{\mu'}_\mu q^{\nu'}_\nu \nabla_{\mu'} n_{\nu'}. \tag{13}$$

This tensor is symmetric and it is connected to the Lie derivative of the intrinsic metric $\mathcal{L}_n q_{\mu\nu} = 2K_{\mu\nu}$. It enters the relation between the Riemann tensor of Σ_t ($\mathcal{R}$) and that of $\mathcal{M}$ (R),

$$\mathcal{R}^\mu_{\nu\rho\sigma} = q^\mu_{\mu'}q^{\nu'}_\nu q^{\rho'}_\rho q^{\sigma'}_\sigma R^{\mu'}_{\nu'\rho'\sigma'} - K_{\nu\sigma}K^\mu_\rho - K_{\nu\rho}K^\mu_\sigma \tag{14}$$

This formula, proved in the Appendix and known as the Gauss-Codazzi equation, lets us rewrite the

action (2) in the following form,[2]

$$S = \int \mathrm{d}t \int_\Sigma \mathrm{d}^3x \sqrt{q} N \left[\mathcal{R} - K^2 + \mathrm{Tr}(KK) \right]. \tag{15}$$

A key consequence of this analysis is the fact that time derivatives of N and N^a do not enter in the Lagrangian L. This implies that N and N^a are Lagrange multipliers, with null conjugate momenta $\delta L/\delta \dot{N} = \delta L/\delta \dot{N}^a = 0$. The true dynamical variables are the spatial components q_{ab} only, with conjugate momenta

$$\pi^{ab} \equiv \frac{\delta L}{\delta \dot{q}_{ab}} = \sqrt{q} \left(K^{ab} - K q^{ab} \right). \tag{16}$$

One can finally evaluate the Legendre transform, which after some calculations gives

$$S_{EH}\left(q_{ab}, \pi^{ab}, N, N^a\right) = \frac{1}{16\pi G} \int \mathrm{d}t \int \mathrm{d}^3x \ \left[\pi^{ab} \dot{q}_{ab} - N^a H_a - N H \right], \tag{17}$$

where (the covariant derivative ∇ contains only spatial indexes)

$$H_a = -2\sqrt{q}\, \nabla_b \left(\frac{\pi^b_a}{\sqrt{q}} \right) \tag{18}$$

and

$$H = \frac{1}{\sqrt{q}} G_{abcd} \pi^{ab} \pi^{cd} - \sqrt{q}\, \mathcal{R}, \qquad G_{abcd} = q_{ac} q_{bd} + q_{ad} q_{bc} - q_{ab} q_{cd}. \tag{19}$$

The quantity G_{abcd} is often called the supermetric, or DeWitt metric. The variation of the action (17) with respect to the Lagrange multipliers gives the equations

$$H_a(q, \pi) = 0, \qquad\qquad H(q, \pi) = 0, \tag{20}$$

which are called respectively vector, or space-diffeomorphism constraint, and scalar or Hamiltonian constraint. In the following we will also use the compact notation $H_\mu = (H, H_a)$. Physical configurations, also called on-shell configurations with particle physics jargon, must satisfy these constraints.

From (17) we see that the Hamiltonian of general relativity is

$$\mathbf{H} = \frac{1}{16\pi G} \int \mathrm{d}^3x \ N^a H_a + N H. \tag{21}$$

This Hamiltonian is peculiar, since it is proportional to the Lagrange multipliers and thus *vanishes* on-shell. Hence, there is no dynamics and no physical evolution in the time t. This puzzling absence of a physical Hamiltonian is in fact a consequence of what we discussed earlier: the diffeomorphism-invariance of the theory tells us that t is a mere parameter devoid of an absolute physical meaning, thus there is no physical dynamics in t. This is the root of the *problem of time* in general relativity. For discussions of this problem, see [21] and [22].

2.2.2 Symplectic structure

From now on, we work in units $16\pi G = 1$. The Hamiltonian formulation (17) allows us to study the phase space of general relativity. It is parametrized by the pair (q_{ab}, π^{ab}), with canonical Poisson brackets

$$\left\{ \pi^{ab}(t, x), q_{cd}(t, x') \right\} = \delta^a_{(c} \delta^b_{d)} \delta(x - x'). \tag{22}$$

[2] We are assuming for simplicity of exposition that Σ has no boundary, the case with boundary is treated extensively in [19] and [20].

From there we can evaluate the following brackets among the constraints (for the explicit calculation see in the Appendix)

$$\begin{aligned}\{H_a(x), H_b(y)\} &= H_a(y)\partial_b\delta(x-y) - H_b(x)\partial'_a\delta(x-y)\\ \{H_a(x), H(y)\} &= H(x)\partial_a\delta(x-y)\\ \{H(x), H(y)\} &= H^a(y)\partial_a\delta(x-y) - H^a(y)\partial'_a\delta(x-y)\end{aligned} \tag{23}$$

Notice that the right-hand sides vanish on the constraint surface (20). This means that the Poisson flows generated by the constraints preserve the constraint hypersurface. Constraints with this characteristic are said to be *first class*, as opposed to *second class* constraints whose Poisson brackets do not vanish on-shell. First class constraints generate gauge transformations on the constraint surface (See e.g. [23] for details).

To see what the gauge transformations look like in our case, consider the smearing of the constraints

$$H(\vec{N}) = \int_\Sigma H^a(x)N_a(x)\mathrm{d}^3x, \qquad H(N) = \int_\Sigma H(x)N(x)\mathrm{d}^3x. \tag{24}$$

An explicit computation (see Appendix) shows that

$$\Big\{H(\vec{N}), q_{ab}\Big\} = \mathcal{L}_{\vec{N}}q_{ab}, \qquad \Big\{H(\vec{N}), \pi^{ab}\Big\} = \mathcal{L}_{\vec{N}}\pi^{ab} \tag{25}$$

which means that the vector constraint is the generator of space-diffeomorphism on Σ. The situation is somewhat subtler for the Hamiltonian constraint. We now have (see Appendix)

$$\{H(N), q_{ab}\} = \mathcal{L}_{\vec{n}N}q_{ab}, \tag{26}$$

$$\Big\{H(N), \pi^{ab}\Big\} = \mathcal{L}_{\vec{n}N}\pi^{ab} + \frac{1}{2}q^{ab}NH - 2N\sqrt{q}q^{c[a}q^{b]d}R_{cd} \tag{27}$$

The first bracket is the action of time diffeomorphisms on q_{ab}. The second bracket gives the action of time diffeomorphisms on π_{ab}, but contains also two extra pieces. These vanish if $H = 0$ and $R_{cd} = 0$, namely on the constraint surface and for physical solutions (recall that in vacuum Einstein's equations read $R_{\mu\nu} = 0$). Therefore, we conclude that the constraints H^μ are the generators of the spacetime diffeomorphism group $\mathrm{Diff}(\mathcal{M})$ on physical configurations.

For general configurations, (23) defines the algebra of hypersurface deformations, often called Dirac algebra or Bargmann-Komar algebra. A characteristic of this algebra is that it is not a Lie algebra. In fact, let us look at the smeared Poisson bracket

$$\{H(N_1), H(N_2)\} = H\left(g^{ab}(N_1\partial_bN_2 - N_2\partial_bN_1)\right). \tag{28}$$

We see that outside the constraint surface, the "structure constants" on the right-hand side contain the field g^{ab} itself. Hence they are not constants at all, and the algebra (23) is not a Lie algebra, unlike $\mathrm{Diff}(\mathcal{M})$. Instead, (23) shows that when we introduce a foliation to define the canonical formalism, we still have the symmetry of diffeomorphisms, which acts changing the foliation, and this new one, which acts deforming it. The two symmetries coincide on physical configurations.

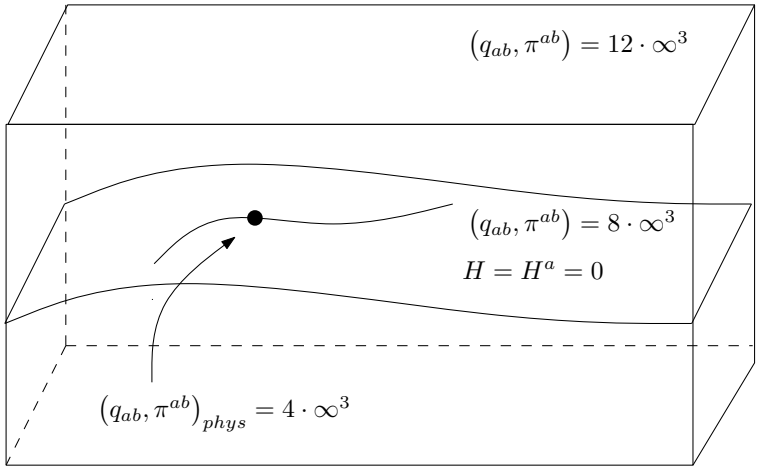
2.2.3 Constraints and physical degrees of freedom

A simple counting of the number of degrees of freedom of general relativity can be done in the covariant perturbative approach described in Section 2.1. The structure of the linearized field equations shows that only two components of $h_{\mu\nu}$ propagate, which correspond to the two helicities of a massless spin 2

particle.

An advantage of the canonical formalism is that it allows us to confirm this counting in a more general and robust way. Recall in fact that in classical physics each point in phase space (i.e. "initial position and momentum") characterizes a physical trajectory, and the number of degrees of freedom is defined to be half the dimensionality of the phase space. In constrained theories, such as general relativity but also gauge theories, one has to be careful with the constraints. To that end, it is customary to distinguish a notion of *kinematical* phase space, and *physical* phase space. The kinematical phase space is the one defined by the Poisson structure of the theory. In our case, the space $\left(q_{ab}, \pi^{ab}\right)$, with Poisson brackets (22). The dimension of this space is $(6+6)\cdot\infty^3 = 12\cdot\infty^3$.

On this space, the constraints define a hypersurface where they are satisfied, i.e. the space of $\left(q_{ab}, \pi^{ab}\right)$ such that $H^\mu(q,\pi) \equiv 0$. We call this the constraint surface. Its dimensionality is $(12-4)\cdot\infty^3 = 8\cdot\infty^3$. The fact that the algebra of constraints is first class guarantees that the gauge transformations generated by the constraints preserve the constraint surface. We shall refer to the trajectories drawn by the gauge transformations as orbits. Points along one orbit correspond to the *same* physical configuration, only described in different coordinate systems. Hence, to select the physical degrees of freedom we have to divide by the gauge orbits, in a manner identical to what happens in gauge theories. Since the orbits span a dimension 4 manifold at each space point, dividing by the orbits gives $(8-4)\cdot\infty^3 = 4\cdot\infty^3$. This is the physical phase space. It has four dimensions per space point, namely the theory has two physical degrees of freedom per space point (or simply two degrees of freedom, for brevity, with the space dependence tacitly implied), a result consistent with the linearized analysis. See e.g. [23] or [25] for more details.



This has far as the counting goes. However, in the case of the linearized analysis we are also able to identify the 2 degrees of freedom as the two helicities, and associate a physical trajectory to each point in phase space, thanks to the fact that we are able to solve the dynamics. Compare this to the mechanics of a point particle: the dynamics is given by $\phi_{[\phi_0,p_0]}(t)$, and the degree of freedom corresponds to varying the initial conditions (ϕ_0, p_0), spanning in this way all possible physical evolutions.

Therefore, if we want to know what the two physical degrees of freedom of general relativity are, we need to control the general solution of the theory. This is a formidable task due to the high non-linearity of the equations, and in spite of the effort in this direction, still little is known. See [26] for a review of some attempts.

There is a particular reformulation of the dynamics that plays a role in quantization. Let us go back to the Gauss-Codazzi equation (14). A simple manipulation gives the so-called first Gauss-Codazzi

identity,

$$\mathcal{R} + K^2 - \mathrm{Tr}(KK) = 2n^\mu n^\nu G_{\mu\nu}, \tag{29}$$

where G is the Einstein tensor. One can show (but we will not do it here) that the left hand side of (29) is equivalent to a linear combination of the constraints. This means that if $g_{\mu\nu}$ satisfies the constraints $H = H^a = 0$ on *any spacelike Cauchy hypersurface*, than it also satisfy all ten Einstein equations $G_{\mu\nu} = 0$. This is a crucial point: *the whole dynamical content of general relativity is in those four constraints.*

2.2.4 Dirac's quantization program

The fact that the constraints include the whole dynamics is at the heart of the theory. It is heavily exploited by the approach to quantization proposed by Dirac, which is based on a definition of dynamical physical states as the ones annihilated by the constraints. The procedure can be schematically divided in three steps:

(i) Find a representation of the phase space variables of the theory as operators in an auxiliary "kinematical" Hilbert space $\mathcal{H}_{kin}$, satisfying the standard commutations relations

$$\{\cdot,\cdot\} \longrightarrow \frac{1}{i\hbar}\,[\cdot,\cdot]\,; \tag{30}$$

(ii) Promote the constraints to operators $\hat{H}^\mu$ in $\mathcal{H}_{kin}$;

(iii) Characterize the space of solutions of the constraints $\mathcal{H}_{phys}$,

$$\hat{H}^\mu\psi = 0 \qquad \forall\psi \in \mathcal{H}_{phys}. \tag{31}$$

These steps should then be completed with an explicit knowledge of the scalar product in $\mathcal{H}_{phys}$ and a physical interpretation of the quantum observables.

Dirac's procedure is more general than gravity: it applies to any fully constrained system. Let us now try to apply it to the ADM formulation of general relativity. We then look for a space of functionals carrying a representation of the quantum Poisson algebra

$$\begin{aligned} \big[\hat{q}_{ab}(x), \hat{\pi}^{cd}(y)\big] &= i\hbar\delta^{cd}_{(ab)}\delta^3(x,y), \\ [\hat{q}_{ab}(x), \hat{g}_{cd}(y)] &= 0, \\ \big[\hat{\pi}^{ab}(x), \hat{\pi}^{cd}(y)\big] &= 0. \end{aligned} \tag{32}$$

Formally, we can proceed by analogy with better known cases, such as a scalar field theory, and consider a Schrödinger representation $\hat{q}_{ab}(x) = q_{ab}(x)$, $\hat{\pi}^{ab}(x) = -i\hbar\frac{\delta}{\delta q_{ab}(x)}$, acting on wave functionals

$$\psi[q_{ab}(x)] \tag{33}$$

of the 3-metric. This procedure, which works well for the scalar field [15], encounters a number of difficulties when applied to the gravitational context. For instance, to define the (kinematical) Hilbert space we need a scalar product, formally

$$\int \mathrm{d}g\, \overline{\psi[g]}\, \psi'[g] \equiv \langle\psi \mid \psi'\rangle\,. \tag{34}$$

However, there is no Lebesgue measure on the space of metrics modulo diffeomorphisms that we can use to define $\mathrm{d}g$. Without this, we can not even check that $\hat{q}_{ab}(x)$ and $\hat{\pi}^{ab}(x)$ are hermitian, nor that $\hat{q}_{ab}(x)$ has positive definite spectrum, as needed to be a spacelike metric.

Let us ignore these and other similar issues, and try to proceed formally assuming that a well-defined $\mathcal{H}_{kin}$ exist. The next step is to promote the constraints (20) to operators, and characterize their space of solutions. Let us proceed in two halves,

$$\mathcal{H}_{kin} \quad \xrightarrow{\hat{H}^a = 0} \quad \mathcal{H}_{Diff} \quad \xrightarrow{\hat{H} = 0} \quad \mathcal{H}_{phys}. \tag{35}$$

Consider first the vector constraint. In the Schrödinger representation defined above, the smeared version gives

$$\hat{H}(N_a)\psi[q_{ab}] = 2i\hbar \int_\Sigma \mathrm{d}^3x \nabla_b N_a \frac{\delta\psi}{\delta q_{ab}} = 0, \tag{36}$$

after an integration by parts. This implies straightforwardly that

$$\psi[q_{ab} + 2\nabla_{(a}N_{b)}] \equiv \psi[q_{ab}],$$

namely the solution of the vector constraint are those functionals of the metric invariants under diffeomorphism. This is very nice, as it realizes at the quantum level the correct action of the classical constraints. However, the space of solutions $\mathcal{H}_{Diff}$ is again ill-defined, since it inherits from the kinematical one the lack of measure theory or other means of control over it.

For the Hamiltonian constraint we can write

$$\hat{H}\psi[q_{ab}] = \left[-\frac{\hbar^2}{2} G^{abcd} : \frac{1}{\sqrt{\det\hat{g}}} \frac{\delta^2}{\delta q_{ab}(x)\delta q_{cd}(x)} : -\sqrt{\det\hat{g}} R(\hat{g}) \right] \psi[q_{ab}], \tag{37}$$

where the colon : means that an ordering of the operators needs to be prescribed. The situation is more complicated for the Hamiltonian constraint, since it requires the definition of products of operators at the same point, notoriously very singular objects. The formal expression (37) is known as the Wheeler-DeWitt equation. Even if manage to give a suitable ordering prescription and regularize the differential operator, the problem with the equation is that we do not have any characterization of the solutions, not even formally as for the diffeo-constraints above. And of course, again no clue on the knowledge of the physical Hilbert space and scalar product (see however minisuperspace models [4]).

Loop quantum gravity is an approach to the problem which improves significantly the situation, and gives a number of answers to these open questions. The key to LQG and to such improvement is surprisingly simple: instead of changing the gravitational theory or the quantization paradigm, we just use different variables to describe gravity. After all, we are familiar with the fact that not all choices of fundamentals variables work out as well when quantizing a classical theory. Consider for instance the harmonic oscillator. Classically, the most elegant description of the system is in terms of action-angle variables, which parametrize the phase space as $\{\phi, J\} = 1$ instead of $\{q, p\} = 1$. However, it is more convenient to quantize the system using the (q, p) variables than the action-angle ones, which require extra care in constructing the operator algebra and dealing with unitarity.

We now introduce the variables that allows us to reformulate general relativity in a way more amenable to Dirac's quantization procedure.

2.3 Appendix

In this Appendix we report some calculations which are used in the main text.

Proof of the Gauss-Codazzi equation. We start the computation from the characterization of $\mathcal{R}$ in terms of covariant derivatives $\tilde{\nabla}$ defined as the projection of ∇ on Σ via q^{μ}_{ν}

$$2\tilde{\nabla}_{[\mu}\tilde{\nabla}_{\nu]}w_{\rho} = -w_{\rho'}\mathcal{R}^{\rho'}_{\rho\mu\nu} \tag{38}$$

$$\tilde{\nabla}_{\mu}\tilde{\nabla}_{\nu}w_{\rho} = \tilde{\nabla}_{\mu}\left(q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\nabla_{\nu'}w_{\rho'}\right) = q^{\mu'}_{\mu}\, q^{\nu''}_{\nu}\, q^{\rho''}_{\rho}\nabla_{\mu'}\left(q^{\nu'}_{\nu''}q^{\rho'}_{\rho''}\nabla_{\nu'}w_{\rho'}\right) \tag{39}$$

it's useful to notice that

$$\begin{aligned} q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\nabla_{\mu'}q^{\rho}_{\nu'} &= q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\nabla_{\mu'}\left(g^{\rho}_{\nu'} + n^{\rho}n_{\nu'}\right) = && (40)\\ &= q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\nabla_{\mu'}g^{\rho}_{\nu'} + q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, n_{\nu'}\nabla_{\mu'}n^{\rho} + q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, n^{\rho}\nabla_{\mu'}n_{\nu'} = && (41)\\ &= n^{\rho}q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\nabla_{\mu'}n_{\nu'} = K_{\mu\nu}n^{\rho} && (42) \end{aligned}$$

we used the definition of extrinsic curvature, orthogonality ($q^{\nu'}_{\nu}\, n_{\nu'} = 0$) and metric compatibility ($\nabla_{\mu'}g^{\rho}_{\nu'} = 0$).

$$\tilde{\nabla}_{\mu}\tilde{\nabla}_{\nu}w_{\rho} = q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\nabla_{\mu'}\nabla_{\nu'}w_{\rho'} + q^{\rho'}_{\rho}\, n^{\nu'}K_{\mu\nu}\nabla_{\nu'}w_{\rho'} + q^{\nu'}_{\nu}K_{\mu\rho}n^{\rho'}\nabla_{\nu'}w_{\rho'} \tag{43}$$

in particular if we consider $w \in \Sigma$ thus $n^{\rho'}w_{\rho'} = 0$

$$n^{\rho'}\nabla_{\nu'}w_{\rho'} = \nabla_{\nu'}n^{\rho'}w_{\rho'} - w_{\rho'}\nabla_{\nu'}n^{\rho'} = -w_{\rho'}\nabla_{\nu'}n^{\rho'} \tag{44}$$

Antisymmetrizing in a,b

$$\begin{aligned} \tilde{\nabla}_{[\mu}\tilde{\nabla}_{\nu]}w_{\rho} &= q^{\mu'}_{[\mu}q^{\nu'}_{\nu]}\, q^{\rho'}_{\rho}\nabla_{\mu'}\nabla_{\nu'}w_{\rho'} + q^{\rho'}_{\rho}\, n^{\nu'}K_{[\mu\nu]}\nabla_{\nu'}w_{\rho'} - q^{\nu'}_{[\nu}K_{\mu]\rho}w_{\rho'}\nabla_{\nu'}n^{\rho'} = && (45)\\ &= q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\nabla_{[\mu'}\nabla_{\nu']}w_{\rho'} - K^{\rho'}_{[\nu}K_{\mu]\rho}w_{\rho'} = && (46)\\ &= -\frac{1}{2}q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\, R^{\rho''}_{\rho'\mu'\nu'}w_{\rho''} - K^{\rho'}_{[\nu}K_{\mu]\rho}w_{\rho'} && (47) \end{aligned}$$

or in terms of the Riemann tensor using (38)

$$\mathcal{R}^{\rho'}_{\rho\mu\nu}w_{\rho'} = \left(q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho''}_{\rho}\, R^{\rho'}_{\rho''\mu'\nu'} + 2K^{\rho'}_{[\nu}K_{\mu]\rho}\right)w_{\rho'} \tag{48}$$

this equation state for $w \in \Sigma$ for make it true for all $v \in \mathcal{M}$ it is sufficient to take $w_{\mu} = q^{\mu'}_{\mu}v_{\mu'}$

$$\mathcal{R}^{\sigma}_{\rho\mu\nu} = q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\, q^{\sigma}_{\sigma'}R^{\sigma'}_{\rho'\mu'\nu'} + 2K^{\sigma}_{[\nu}K_{\mu]\rho} \tag{49}$$

Or equivalently

$$\mathcal{R}_{\sigma\rho\mu\nu} = q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\, q^{\sigma'}_{\sigma}\, R_{\sigma'\rho'\mu'\nu'} + 2K_{\sigma[\nu}K_{\mu]\rho} \tag{50}$$

$$\begin{aligned} \mathcal{R} &= q^{\sigma\mu}q^{\nu\rho}q^{\mu'}_{\mu}\, q^{\nu'}_{\nu}\, q^{\rho'}_{\rho}\, q^{\sigma'}_{\sigma}\, R_{\sigma'\rho'\mu'\nu'} + q^{\sigma\mu}q^{\nu\rho}K_{\sigma\nu}K_{\mu\rho} - q^{\sigma\mu}q^{\nu\rho}K_{\sigma\mu}K_{\nu\rho} = && (51)\\ &= q^{\sigma'\mu'}q^{\nu'\rho'}R_{\sigma'\rho'\mu'\nu'} + \mathrm{Tr}(KK) - K^2 = && (52)\\ &= R + 2n^{\nu'}n^{\rho'}R_{\rho'\nu'} + n^{\sigma'}n^{\mu'}n^{\nu'}n^{\rho'}R_{\sigma'\rho'\mu'\nu'} + \mathrm{Tr}(KK) - K^2 && (53) \end{aligned}$$

the term $n^{\sigma'}n^{\mu'}n^{\nu'}n^{\rho'}R_{\sigma'\rho'\mu'\nu'}$ vanishes because of the symmetries of the Riemann tensor, using the fact that $n^{\nu'}n^{\rho'}g_{\rho'\nu'} = -1$ we can write

$$\mathcal{R} = 2n^{\nu'}n^{\rho'}\left(R_{\rho'\nu'} - \frac{g_{\rho'\nu'}}{2}R\right) + \mathrm{Tr}(KK) - K^2 \tag{54}$$

That is the first Gauss-Codazzi equation.

Constraint Algebra. We define the Poisson bracket as

$$\{A(y), B(z)\} = \int \mathrm{d}^3x \frac{\delta A(y)}{\delta q_{ab}(x)} \frac{\delta B(z)}{\delta \pi^{ab}(x)} - \frac{\delta B(z)}{\delta q_{ab}(x)} \frac{\delta A(y)}{\delta \pi^{ab}(x)} \tag{55}$$

We start the computation from the diffeomorphism constraint

$$H(\vec{N}) = \int \mathrm{d}^3x N^l H_l \tag{56}$$

we want to show that it generates infinitesimal spatial diffeomorphism (it's sufficient to do it on the phase space variables)

$$\left\{q_{ab}(y), H(\vec{N})\right\} = \int \mathrm{d}^3x N^l \frac{\delta H_l(x)}{\delta \pi^{ab}(y)} = 2\nabla_{(a} N_{b)} = L_{\vec{N}} q_{ab} \tag{57}$$

in fact from the definition of the Lie derivatives

$$\mathcal{L}_{\vec{N}} q_{ab} = (\nabla_a N^c) q_{cb} + (\nabla_b N^c) q_{ac} - \underbrace{N^c \nabla_c q_{ab}}_{=0} = 2\nabla_{(a} N_{b)}$$

Note that the functional derivative is easy to compute with a by part integration. The last term vanishes for metric compatibility and in the Lie derivative you can substitute ∂ with ∇.

$$\begin{aligned}
\left\{\pi^{ab}, H(\vec{N})\right\} &= \int -N^l \frac{\delta H_l(x)}{\delta q_{ab}(y)} = -2\pi^{ca}\nabla_c N^b + \sqrt{q}\nabla_k \left(\frac{N^k}{\sqrt{q}}\pi^{ab}\right) = \\
&= -2\pi^{ca}\nabla_c N^b \frac{\sqrt{q}}{\sqrt{q}} + \sqrt{q}(\nabla_k N^k)\frac{\pi^{ab}}{\sqrt{q}} + \sqrt{q}N^k \nabla_k \frac{\pi^{ab}}{\sqrt{q}} = \\
&= \sqrt{q}\mathcal{L}_{\vec{N}} \frac{\pi^{ab}}{\sqrt{q}} + \frac{\pi^{ab}}{\sqrt{q}}(\nabla_k N^k)\sqrt{q} = \\
&= \sqrt{q}\mathcal{L}_{\vec{N}} \frac{\pi^{ab}}{\sqrt{q}} + \frac{\pi^{ab}}{\sqrt{q}}\mathcal{L}_{\vec{N}}\sqrt{q} \equiv \mathcal{L}_{\vec{N}}\pi^{ab}
\end{aligned}$$

Note that π^{ab} is a tensorial density so its lie derivative has to be defined as above. Given A an arbitrary function of q and π using Leibniz we can compute the Poisson bracket easily

$$\begin{aligned}
\left\{A, H(\vec{N})\right\} &= \int \frac{\delta A}{\delta q_{ab}} \frac{\delta H(\vec{N})}{\delta \pi^{ab}} - \frac{\delta H(\vec{N})}{\delta q_{ab}} \frac{\delta A}{\delta \pi^{ab}} = \\
&= \int \frac{\delta A}{\delta q_{ab}} \left\{q_{ab}, H(\vec{N})\right\} - \left\{\pi^{ab}, H(\vec{N})\right\} \frac{\delta A}{\delta \pi^{ab}} = \\
&= \int \frac{\delta A}{\delta q_{ab}} \mathcal{L}_{\vec{N}} q_{ab} - \mathcal{L}_{\vec{N}}\pi^{ab} \frac{\delta A}{\delta \pi^{ab}} = \mathcal{L}_{\vec{N}} A
\end{aligned}$$

In particular for H and H_a as function of h and π we deduce

$$\begin{aligned}
\mathcal{L}_{\vec{N}} H_b &= \sqrt{q}\mathcal{L}_{\vec{N}} \frac{H_b}{\sqrt{q}} + \frac{H_b}{\sqrt{q}}\mathcal{L}_{\vec{N}}\sqrt{q} = \\
&= \sqrt{q}N^a \partial_a \frac{H_b}{\sqrt{q}} + H_a \partial_b N^a + H_b \nabla_a N^a = \\
&= N^a \partial_a H_b + N^a H_b \sqrt{q}\partial_a \frac{1}{\sqrt{q}} + H_a \partial_b N^a + H_b \nabla_a N^a = \\
&= N^a \partial_a H_b - N^a H_b \partial_a(\sqrt{q})\frac{1}{\sqrt{q}} + H_a \partial_b N^a + H_b \nabla_a N^a = \\
&= N^a \partial_a H_b + H_a \partial_b N^a + H_b(\nabla_a N^a - \frac{N^a}{\sqrt{q}}\partial_a \sqrt{q}) = \\
&= N^a \partial_a H_b + H_a \partial_b N^a + H_b \partial_a N^a
\end{aligned}$$

In fact

$$\nabla_a N^a = \partial_a N^a + \Gamma^a_{ba} N^b$$
$$\frac{N^a}{\sqrt{q}} \partial_a \sqrt{q} = \frac{1}{2} N^a q^{ab} \partial_a q_{ab} = \Gamma^a_{ba} N^b$$

Where the last equality is due to metric compatibility. We can conclude that

$$\left\{ \int N^a H_a, H_b \right\} = -N^a \partial_a H_b - H_a \partial_b N^a - H_b \partial_a N^a \tag{58}$$

That is the smeared version of

$$\{H_a(x), H_b(y)\} = H_a(y) \partial_b \delta(x-y) - H_b(x) \partial'_a \delta(x-y) \tag{59}$$

Moreover also the second Poisson bracket is easy to compute using the fact that $H(\vec{N})$ is the generator of infinitesimal diffeomorphism along $\vec{N}$.

$$\begin{aligned}
\mathcal{L}_{\vec{N}} H &= \sqrt{q} \mathcal{L}_{\vec{N}} \frac{H}{\sqrt{q}} + \frac{H}{\sqrt{q}} \mathcal{L}_{\vec{N}} \sqrt{q} = \\
&= \sqrt{q} N^a \partial_a \frac{H}{\sqrt{q}} + \frac{H}{\sqrt{q}} (\frac{1}{2} \sqrt{q} q^{cd} \mathcal{L}_{\vec{N}} q_{cd}) = \\
&= N^a \partial_a H + \sqrt{q} N^a H \partial_a \frac{1}{\sqrt{q}} + \frac{H}{\sqrt{q}} \frac{2\sqrt{q}}{2} q^{cd} q_{ca} \nabla_d N^a = \\
&= N^a \partial_a H + H \partial_a N^a + \sqrt{q} N^a H (\frac{-1}{2}) \frac{1}{\sqrt{q}} q^{cd} \partial_a q_{cd} + H \Gamma^a_{ba} N^b = \\
&= N^a \partial_a H + H \partial_a N^a + \sqrt{q} N^a H (\frac{-1}{2}) \frac{1}{\sqrt{q}} q^{cd} \underbrace{\nabla_a q_{cd}}_{=0} = N^a \partial_a H + H \partial_a N^a
\end{aligned} \tag{60}$$

putting all together

$$\left\{ \int N^a H_a, H \right\} = -N^a \partial_a H - H \partial_a N^a \tag{61}$$

that is equivalent to

$$\{H_a(x), H(y)\} = H \partial_a \delta(x-y) \tag{62}$$

To ultimate the computation of the second Poisson bracket we need some preliminary results. Given an arbitrary function $f(\pi, h)$

$$\left\{ \left\{ f, H(\vec{N}_1) \right\}, H(\vec{N}_2) \right\} = \mathcal{L}_{\vec{N}_1} \mathcal{L}_{\vec{N}_2} f \tag{63}$$

using the Jacobi identity

$$\begin{aligned}
\left\{ f, \left\{ H(\vec{N}_1), H(\vec{N}_2) \right\} \right\} &= \left\{ \left\{ H(\vec{N}_2), f \right\}, H(\vec{N}_1) \right\} + \left\{ \left\{ f, H(\vec{N}_1) \right\}, H(\vec{N}_2) \right\} = \\
&= \left(-\mathcal{L}_{\vec{N}_1} \mathcal{L}_{\vec{N}_2} + \mathcal{L}_{\vec{N}_2} \mathcal{L}_{\vec{N}_1} \right) f = \left[\mathcal{L}_{\vec{N}_2}, \mathcal{L}_{\vec{N}_1} \right] f = \\
&= \mathcal{L}_{[\vec{N}_2, \vec{N}_1]} f = \left\{ f, \int [N_2, N_1]^a H_a \right\}
\end{aligned}$$

If we smear the (61) we obtain:

$$\left\{ \int N_1^a H_a, \int N H \right\} = \int \mathcal{L}_{\vec{N}} N H \tag{64}$$

$$\left\{ \int N_1^a H_a, \int N_2^a H_a \right\} = \int [N_1, N_2]^a H_a \tag{65}$$

Finally we compute the last Poisson bracket, if $N \neq N'$ (if they are equal it vanish)

$$\left\{ H(N), H(N') \right\} = \int \mathrm{d}^3 x \frac{\delta H(N)}{\delta q_{ab}(x)} \frac{\delta H(N')}{\delta \pi^{ab}(x)} - \frac{\delta H(N')}{\delta q_{ab}(x)} \frac{\delta H(N)}{\delta \pi^{ab}(x)} \tag{66}$$

We notice that all the algebraical terms in q and π simplify each other. In fact

$$\left.\frac{\delta H(x)}{\delta q_{ab}(y)}\right|_{\substack{algebraic\\part}} = f^h(x)_{ab}\delta(x-y) \tag{67}$$

$$\left.\frac{\delta H(x)}{\delta \pi_{ab}(y)}\right|_{\substack{algebraic\\part}} = f_\pi(x)^{ab}\delta(x-y) \tag{68}$$

$$\left\{ H(N)|_{\substack{algebraic\\part}}, H(N')\big|_{\substack{algebraic\\part}} \right\} = \int \mathrm{d}^3x N f^h_{ab} N' f^{ab}_\pi - N' f^{ab}_\pi N f^h_{ab} = 0 \tag{69}$$

The non algebraic part in q are in R conversely the Hamiltonian constraint is algebraic in π.

$$\{H(N), H(N')\} = \int \mathrm{d}^3x \frac{\delta H(N)^{\substack{non\\algebraic\\term}}}{\delta q_{ab}(x)} \frac{\delta H(N')}{\delta \pi^{ab}(x)} - \frac{\delta H(N')^{\substack{non\\algebraic\\term}}}{\delta q_{ab}(x)} \frac{\delta H(N)}{\delta \pi^{ab}(x)} \tag{70}$$

We compute (ignoring algebraic terms for simplicity)

$$\begin{aligned}
\delta_q \int NH &= -\int N\sqrt{q} q^{ab} \delta R_{ab} = \\
&= -\int N\sqrt{q}\left(\nabla_a\nabla_b\delta q_{cd} q^{ac} q^{bd} - q^{cd}\nabla^b\nabla_b\delta q_{cd}\right) = \\
&= -\int \delta q_{cd}\sqrt{q}\left(\nabla^c\nabla^d N - \nabla^2(Nq^{cd})\right)
\end{aligned}$$

Substituting this variation in the Poisson bracket the first term becomes

$$\begin{aligned}
\int \frac{\delta\int NH}{\delta q_{ab}} \frac{\delta \int N'H}{\delta \pi^{ab}} &= -\int \sqrt{q}\left(\nabla^a\nabla^b N - \nabla^2(Nq^{ab})\right) 2N'\left(\frac{\pi_{ab}}{\sqrt{q}} - \frac{\pi q_{ab}}{\sqrt{q}(3-1)}\right) = \\
&= -2\int (\nabla^a\nabla^b N)\pi_{ab}N' - (\nabla^2 N)\pi N' - (\nabla^2 N)N'\frac{\pi}{3-1} + (\pi\nabla^2 N)\frac{q^{ab}q_{ab}}{3-1}N' = \\
&= 2\int (\nabla^a\nabla^b N)\pi_{ab}N' + \left(\frac{3}{3-1} - \frac{1}{3-1} - 1\right)\pi N'\nabla^2 N = \\
&= -2\int \sqrt{q}(\nabla^a\nabla^b N)\frac{\pi_{ab}}{\sqrt{q}}N' = 2\int \sqrt{q}\nabla^b N \nabla^a\left(\frac{\pi_{ab}N'}{\sqrt{q}}\right)
\end{aligned} \tag{71}$$

summing also the second term

$$\begin{aligned}
[70] &= 2\int \sqrt{q}\left[\nabla^b N\nabla^a\left(\frac{\pi_{ab}N'}{\sqrt{q}}\right) - \nabla^b N'\nabla^a\left(\frac{\pi_{ab}N}{\sqrt{q}}\right)\right] = \\
&= 2\int \sqrt{q}\left[\nabla^b N\nabla^a N'\frac{\pi_{ab}}{\sqrt{q}} + N'(\nabla^b N)\nabla^a\left(\frac{\pi_{ab}}{\sqrt{q}}\right) - \nabla^b N'\nabla^a N\frac{\pi_{ab}}{\sqrt{q}} + N(\nabla^b N')\nabla^a\left(\frac{\pi_{ab}}{\sqrt{q}}\right)\right] = \\
&= \int \left(-N'\nabla^b N + N\nabla^b N'\right) 2\sqrt{q}\nabla^a\left(\frac{\pi_{ab}}{\sqrt{q}}\right) = \int \left(N\nabla^b N' - N'\nabla^b N\right) H_a
\end{aligned} \tag{72}$$

3 Tetrad formulation

A tetrad is a quadruple of 1-forms, $e^I_\mu(x)$, $I = 0, 1, 2, 3$ such that

$$g_{\mu\nu}(x) = e^I_\mu(x) e^J_\nu(x) \eta_{IJ}. \tag{73}$$

By its definition, it provides a local isomorphism between a general reference frame and an inertial one, characterized by the flat metric η_{IJ}. A local inertial frame is defined up to a Lorentz transformation, and in fact notice that the definition is invariant under

$$e^I_\mu(x) \longrightarrow \tilde{e}^I_\mu(x) = \Lambda^I_J(x) e^J_\mu(x). \tag{74}$$

This means that the "internal" index I carries a representation of the Lorentz group. Contracting vectors and tensors in spacetime with the tetrad, we get objects that transform under the Lorentz group, e.g. $e^I_\mu n^\mu = n^I$. The tetrad thus provides an isomorphism between the tangent bundle of $\mathcal{M}$, $T(\mathcal{M}) = \bigcup_\rho (\rho, T_\rho(\mathcal{M}))$, and a Lorentz principal bundle $F = (\mathcal{M}, SO(3,1))$. On this bundle we have a connection ω^{IJ}_μ, that is a 1-form with values in the Lorentz algebra, which we can use to define covariant differentiation of the fibers,

$$D_\mu v^I(x) = \partial_\mu v^I(x) + \omega^I_{\mu J}(x) v^J(x). \tag{75}$$

This is the analogue of the covariant derivative $\nabla_\mu = \partial_\mu + \Gamma_\mu$ for vectors in $T(\mathcal{M})$. We can also define the derivative for objects with both indices, such as the tetrad,

$$\mathcal{D}_\mu e^I_\nu = \partial_\mu e^I_\nu + \omega^I_{\mu J} e^J_\nu - \Gamma^\rho_{\nu\mu} e^I_\rho. \tag{76}$$

As the Levi-Civita connection $\Gamma(g)$ is metric-compatible, i.e. $\nabla_\mu g_{\nu\rho} = 0$, we require ω_μ to be tetrad-compatible, i.e. $\mathcal{D}_\mu e^I_\nu \equiv 0$, and call it spin connection. This implies

$$\partial_{(\mu} e^I_{\nu)} + \omega^I_{(\mu J} e^J_{\nu)} = \Gamma^\rho_{(\nu\mu)} e^I_\rho, \qquad \partial_{[\mu} e^I_{\nu]} + \omega^I_{[\mu J} e^J_{\nu]} = \Gamma^\rho_{[\nu\mu]} e^I_\rho \equiv 0, \tag{77}$$

where we separated the spacetime indices into their symmetric and antisymmetric combinations, and used the fact that the Levi-Civita connection $\Gamma(g)$ has no antisymmetric part.

From these equations we immediately obtain the following relation between the spin and Levi-Civita connections,

$$\omega^I_{\mu J} = e^I_\nu \nabla_\mu e^\nu_J, \tag{78}$$

as well as the fact that the spin connection satisfies

$$\mathrm{d}_\omega e^I = \mathrm{d}e^I + \omega^I{}_J \wedge e^J = \left(\partial_\mu e^I_\nu + \omega^I_{\mu J} e^J_\nu\right) \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = 0. \tag{79}$$

This equation is known as Cartan's first structure equation. Here we introduced the exterior calculus of forms, with d the exterior derivative, d_ω the covariant exterior derivative, and $\wedge$ the wedge product. See [27] for an introduction to this formalism.

Given the connection, we define its curvature

$$F^{IJ} = \mathrm{d}\omega^{IJ} + \omega^I{}_K \wedge \omega^{KJ}, \tag{80}$$

whose components are

$$F^{IJ}_{\mu\nu} = \partial_\mu \omega^{IJ}_\nu - \partial_\nu \omega^{IJ}_\mu + \omega^I{}_{K\mu} \omega^{KJ}_\nu - \omega^J{}_{K\mu} \omega^{KI}_\nu. \tag{81}$$

Using the solution $\omega(e)$ given in (78), an explicit calculation gives

$$F_{\mu\nu}^{IJ}\left(\omega\left(e\right)\right) \equiv e^{I\rho}e^{J\sigma}R_{\mu\nu\rho\sigma}(e), \tag{82}$$

where $R_{\mu\nu\rho\sigma}(e)$ is the Riemann tensor constructed out of (the metric defined by) the tetrad e_μ^I. This relation is known as Cartan second structure equation. It shows that general relativity is a gauge theory whose local gauge group is the Lorentz group, and the Riemann tensor is nothing but the field-strength of the spin connection.

Cartan second structure equation. Starting from the definition of F in (81) and inserting (78), we have

$$\begin{aligned} F_{\mu\nu}^{IJ} =& \partial_\mu e_\rho^I \partial_\nu e^{\rho J} + \partial_\mu e_\rho^I \Gamma_{\sigma\nu}^\rho e^{\sigma J} + e_\rho^I \partial_\mu \left(\Gamma_{\sigma\nu}^\rho\right) e^{\sigma J} + e_\rho^I \Gamma_{\sigma\nu}^\rho \partial_\mu e^{\sigma J} + e_\rho^I \partial_\mu e_K^\rho e_\sigma^K \partial_\nu e^{\sigma J} + \\ &+ e_\rho^I \Gamma_{\delta\mu}^\rho e_K^\delta e_\sigma^K \partial_\nu e^{\sigma J} + e_\rho^I \partial_\mu e_K^\rho e_\sigma^K \Gamma_{\delta\nu}^\sigma e^{\delta J} + e_\rho^I \Gamma_{\delta\mu}^\rho e_K^\delta e_\sigma^K \Gamma_{\eta\nu}^\sigma e^{\eta J} - (\mu \leftrightarrow \nu). \end{aligned}$$

Next, we use $e_\rho^I e_K^\rho = \delta_K^I$ and $e_\rho^I \partial_\mu e_K^\rho = -\partial_\mu (e_\rho^I) e_K^\rho$ to rewrite this expression as

$$\begin{aligned} F_{\mu\nu}^{IJ} &= e_\rho^I e^{\sigma J} \partial_\mu \left(\Gamma_{\sigma\nu}^\rho\right) + e_\rho^I e^{\sigma J} \Gamma_{\delta\mu}^\rho \Gamma_{\sigma\nu}^\delta + \partial_m u e_\rho^I \partial_\nu e^{\rho J} - \partial_\mu e_\rho^I \partial_\nu e^{\rho J} + \\ &+ \partial_\mu e_\rho^I \Gamma_{\sigma\nu}^\rho e^{\sigma J} - \partial_\mu e_\rho^I \Gamma_{\sigma\nu}^\rho e^{\sigma J} + e_\rho^I \Gamma_{\sigma\nu}^\rho \partial_\mu e^{\sigma J} + e_\rho^I \Gamma_{\sigma\mu}^\rho \partial_\nu e^{\sigma J} - (\mu \leftrightarrow \nu) \\ &= 2 e_\rho^I e^{J\sigma} \left(\partial_{(\mu} \Gamma_{\sigma\nu)}^\rho + \Gamma_{\delta(\mu}^\rho \Gamma_{\sigma\nu)}^\delta\right) = e^{I\rho} e^{\sigma J} R_{\mu\nu\rho\sigma}. \end{aligned}$$

Relation between the determinants. The determinant g of $g_{\mu\nu}$ is related to the determinant of the tetrad e by the simple relation

$$g = -e^2. \tag{83}$$

This expression can be easily derived recalling Caley's formula for the determinant of a matrix,

$$g = \det g_{\mu\nu} = \frac{1}{4!} \varepsilon^{\mu\nu\rho\sigma} \varepsilon^{\alpha\beta\gamma\delta} g_{\mu\alpha} g_{\nu\beta} g_{\rho\gamma} g_{\sigma\delta}. \tag{84}$$

If we substitute the expression of $g_{\mu\nu}$ in terms of tetrads we get

$$\begin{aligned} g &= \det\left(e_\mu^I e_\nu^J \eta_{IJ}\right) = \frac{1}{4!} \varepsilon^{\mu\nu\rho\sigma} \varepsilon^{\alpha\beta\gamma\delta} e_\mu^I e_\alpha^J \eta_{IJ} e_\nu^K e_\beta^L \eta_{KL} e_\rho^M e_\gamma^N \eta_{MN} e_\sigma^O e_\delta^P \eta_{OP} = \\ &= \frac{1}{4!} \varepsilon^{\mu\nu\rho\sigma} e_\mu^I e_\nu^K e_\rho^M e_\sigma^O \varepsilon^{\alpha\beta\gamma\delta} e_\alpha^J e_\beta^L e_\gamma^N e_\delta^P \eta_{IJ} \eta_{KL} \eta_{MN} \eta_{OP} = \\ &= \frac{1}{4!} e^2 \varepsilon^{IKMO} \varepsilon^{JLNP} \eta_{IJ} \eta_{KL} \eta_{MN} \eta_{OP} = -e^2 \end{aligned}$$

3.1 The action in terms of tetrads

The Einstein-Hilbert action can be rewritten as a functional of the tetrad in the following way (recall we take units $16\pi G = 1$),

$$S_{EH}(e_\mu^I) = \frac{1}{2} \varepsilon_{IJKL} \int e^I \wedge e^J \wedge F^{KL}\left(\omega(e)\right). \tag{85}$$

On top of the invariance under diffeomorphism, this reformulation of the theory possesses an additional gauge symmetry under local Lorentz transformations.

Rewriting the actions in terms of tetrads. Explicitly using (82) and (83)

$$\begin{aligned} S_{EH}(g_{\mu\nu}(e)) &= \int \mathrm{d}^4x\,\sqrt{-g}\,g^{\mu\nu}R_{\mu\nu} = \int \mathrm{d}^4x\,e\,e^{\mu}_I e^{\nu I}R_{\mu\rho\nu\sigma}e^{\rho}_J e^{\sigma J} = \\ &= \int \mathrm{d}^4x e e^{\mu}_I e^{\rho}_J F^{IJ}_{\mu\rho}\left(\omega(e)\right) = \int \mathrm{d}^4x \frac{1}{4}\varepsilon_{IJKL}\varepsilon^{\mu\rho\alpha\beta}e^K_\alpha e^L_\beta F^{IJ}_{\mu\rho}\left(\omega(e)\right) \\ &= \int \frac{1}{2}\varepsilon_{IJKL}e^I\wedge e^J\wedge F^{KL}\left(\omega(e)\right) \end{aligned}$$

A fact which plays an important role in the following is that we can lift the connection to be an *independent variable*, and consider the new action

$$S(e^I_\mu,\omega^{IJ}_\mu) = \frac{1}{2}\varepsilon_{IJKL}\int e^I\wedge e^J\wedge F^{KL}\left(\omega\right). \tag{86}$$

Although it depends on extra fields, this action remarkably gives the same equations of motion as the Einstein-Hilbert one (85). This happens because the extra field equations coming from varying the action with respect to ω do not add anything new: they simply impose the form (78) of the spin connection, and general relativity is thus recovered.

Deriving the fields equations. One first verifies the Palatini identity

$$\delta_\omega F^{KL}\left(\omega\right) = \mathrm{d}_\omega\delta\omega^{KL}. \tag{87}$$

Then, the variation gives

$$\delta_\omega S = \frac{1}{2}\varepsilon_{IJKL}\int e^I\wedge e^J\wedge \mathrm{d}_\omega\delta\omega^{KL} = -\frac{1}{2}\varepsilon_{IJKL}\int \mathrm{d}_\omega\left(e^I\wedge e^J\right)\wedge\delta\omega^{KL} \tag{88}$$

after an integration by part. Imposing the vanishing of the variation, we obtain the field equation

$$\varepsilon_{IJKL}e^I\wedge \mathrm{d}_\omega e^J = 0. \tag{89}$$

If the tetrad is invertible this equation implies $d_\omega e^J = 0$, which in turns implies (78) in terms of the Levi-Civita connection of the metric associated with e^I_μ.

As it gives the same field equations, (87) can be used as the action of general relativity. Notice that only first derivatives appears, thus it provides a first order formulation of general relativity. Furthermore, the action is polynomial in the fields, a desirable property for quantization. On the other hand, there are two non-trivial aspects to take into account:

- The equivalence with general relativity holds only if the tetrad is *non-degenerate*, i.e. invertible. On the other hand, (87) is also defined for degenerate tetrads, since inverse tetrads never appear. Compare the situation with the Einstein-Hilbert action, where the inverse metric appears explicitly. Hence, the use of (87) leads naturally to an extension of general relativity where a sector with degenerate tetrads, and thus degenerate metrics, exists.

- If we insist on the connection being an independent variable, there exists a second term that we can add to the Lagrangian that is compatible with all the symmetries and has mass dimension 4:

$$\delta_{IJKL}e^I\wedge e^J\wedge F^{KL}(\omega), \tag{90}$$

  where $\delta_{IJKL}\equiv\delta_{I[K}\delta_{L]J}$. This term is not present in the ordinary second order metric, since when

(78) holds,

$$\delta_{IJKL} e^I \wedge e^J \wedge F^{KL}(\omega(e)) = \epsilon^{\mu\nu\rho\sigma} R_{\mu\nu\rho\sigma}(e) \equiv 0. \tag{91}$$

Adding this second term to (87) with a coupling constant $1/\gamma$ leads to the so-called Holst action [28]

$$S(e,\omega) = \left(\frac{1}{2}\varepsilon_{IJKL} + \frac{1}{\gamma}\delta_{IJKL}\right) \int e^I \wedge e^J \wedge F^{KL}(\omega). \tag{92}$$

Assuming non-degenerate tetrads, this action leads to the same field equations of general relativity,

$$\omega_\mu^{IJ} = e_\nu^I \nabla_\mu e^{J\nu}, \qquad G_{\mu\nu}(e) = 0. \tag{93}$$

This result is *completely independent* of the value of γ, which is thus a parameter irrelevant in classical vacuum general relativity. It will however turn out to play a key role in the quantum theory, where it is known as the Immirzi parameter.[3]

Deriving the fields equations for the Holst action. The vanishing of the variations give

$$\left(\frac{1}{2}\varepsilon_{IJKL} + \frac{1}{\gamma}\delta_{IJKL}\right) e^I \wedge \mathrm{d}_\omega e^J = 0, \tag{94}$$

$$\left(\frac{1}{2}\varepsilon_{IJKL} + \frac{1}{\gamma}\delta_{IJKL}\right) e^J \wedge F^{KL}(\omega) = 0. \tag{95}$$

For invertible tetrads, the first one is again uniquely solved by (78). Substituting this solution into the second equation we get

$$\begin{aligned} &e^{I\nu} G_\nu^\alpha + \frac{1}{\gamma}\delta_{IJKL}\varepsilon^{\mu\sigma\delta\alpha} e_\mu^J F_{\sigma\delta}^{KL} = e^{I\nu} G_\nu^\alpha + \frac{1}{\gamma}\delta_{IJKL}\varepsilon^{\mu\sigma\delta\alpha} e_\mu^J e^{K\nu} e^{L\rho} R_{\sigma\delta\nu\rho} = \\ &= e^{I\nu} G_\nu^\alpha + \frac{1}{2\gamma}\varepsilon^{\mu\sigma\delta\alpha}\left(e^{I\nu}\delta_\mu^\rho - e^{I\rho}\delta_\mu^\nu\right) R_{\sigma\delta\nu\rho} = e^{I\nu} G_\nu^\alpha - \frac{1}{\gamma}\varepsilon^{\sigma\delta\mu\alpha} e^{I\nu} R_{\sigma\delta\mu\nu} = e^{I\nu} G_\nu^\alpha = 0, \end{aligned}$$

which is equivalent to $G_{\mu\nu} = 0$ thanks to the non-degeneracy of the tetrad. In the last step we used the first Bianchi identity $\varepsilon^{\sigma\delta\mu\alpha} R_{\sigma\delta\mu\nu} = 0$.

3.2 Hamiltonian analysis of tetrad formulation

For the Hamiltonian formulation we proceed as before, assuming a $3+1$ splitting of the space-time ($\mathcal{M} \cong \mathbb{R} \times \Sigma$) and coordinates (t, x). We introduce the lapse function and the shift vector (N, N^a) and the ADM decomposition of the metric (11). It is easy to see that a tetrad for the ADM metric is given by

$$e_0^I = e_\mu^I \tau^\mu = N n^I + N^a e_a^I, \qquad \delta_{ij} e_a^i e_b^j = g_{ab},\ i = 1,2,3. \tag{96}$$

The "triad" e_a^i is the spatial part of the tetrad. As before, we want to identify canonically conjugated variables and perform the Legendre transform, but we now have two new features which complicate the analysis. The first one is the tetrad formulation, which in particular has introduced a new symmetry in the action: the invariance under local Lorentz transformations. As a consequence, we expect more constraints to appear, corresponding to the generators of the new local symmetry. The second one is the use of the tetrad and the connection as independent fields. Therefore, the conjugate variables are now functions of both e_a^I and ω_a^{IJ} (and their time derivatives), as opposed to be functions of the metric g_{ab} only.

[3] The Immirzi parameter becomes relevant also at the classical level if source of torsion are present [29, 30].

The consequence of these novelties is a much more complicated structure than in the metric case. In particular, the constraint algebra is *second class*. However, there is a particular choice of variables which simplifies the analysis, making it possible to implement a part of the constraint and reducing the remaining ones to first class again. These are the famous Ashtekar variables, which we now introduce.[4]

To simplify the discussion, it is customary to work in the "time gauge" $e^I_\mu n^\mu = \delta^I_0$, where

$$e^0_\mu = (N, 0) \longrightarrow e^I_0 = \left(N, N^a e^i_a\right). \tag{97}$$

The crucial change of variables is the following: we define the *densitized triad*

$$E^a_i = e e^a_i = \frac{1}{2}\varepsilon_{ijk}\varepsilon^{abc} e^j_b e^k_c, \tag{98}$$

and the *Ashtekar-Barbero connection*

$$A^i_a = \gamma\omega^{0i}_a + \frac{1}{2}\varepsilon^i_{jk}\omega^{jk}_a. \tag{99}$$

These variables turns out to be conjugated. In fact, we can rewrite the action (92) in terms of the new variables as [31, 24]

$$S(A, E, N, N^a) = \frac{1}{\gamma}\int \mathrm{d}t \int_\Sigma \mathrm{d}^3x \left[\dot{A}^i_a E^a_i - A^i_0 D_a E^a_i - NH - N^a H_a\right], \tag{100}$$

where

$$G_j \equiv D_a E^a_i = \partial_a E^a_j + \varepsilon_{jk\ell} A^j_a E^{a\ell}, \tag{101}$$

$$H_a = \frac{1}{\gamma} F^j_{ab} E^b_j - \frac{1+\gamma^2}{\gamma} K^i_a G_i, \tag{102}$$

$$H = \left[F^j_{ab} - \left(\gamma^2 + 1\right)\varepsilon_{jmn} K^m_a K^n_b\right] \frac{\varepsilon_{jk\ell} E^a_k E^b_\ell}{\det E} + \frac{1+\gamma^2}{\gamma} G^i \partial_a \frac{E^a_i}{\det E}. \tag{103}$$

The resulting action is similar to (17), with (A, E) as canonically conjugated variables, as opposed to (q, π). Lapse and shift are still Lagrange multipliers, and consistently we still refer to $H(A, E)$ and $H_a(A, E)$ as the Hamiltonian and space-diffeomorphism constraints.

The algebra is still first class. The new formulation in terms of tetrads has introduced the extra constraint (101). The reader familiar with gauge theories will recognize it as the Gauss constraint. Just as the H^μ constraints generate diffeomorphisms, the Gauss constraint generates gauge transformations. It is in fact easy to check that E^b_j and A^i_a transform respectively as an SU(2) vector and as an SU(2) connection under this transformation.

The algebra generated by the Gauss constraint. We define the smearing of the Gauss constraint

$$G(\Lambda) = \int \mathrm{d}^3x G_i(x)\Lambda^i(x). \tag{104}$$

Its Poisson bracket with the canonical variables give

$$\left\{\int \mathrm{d}^3x \Lambda^j(x) G_j(x), E^a_i(y)\right\} = \int \mathrm{d}^3x \Lambda^j(x)\left\{\partial_b E^b_j + \varepsilon_{mjn} A^m_b E^{bn}, E^a_i(y)\right\} =$$
$$= \gamma\varepsilon_{mjn}\Lambda^j(y) E^{bn}(y)\delta^a_b\delta^m_i = \gamma\varepsilon_{ijn}\Lambda^j(y)E^{an}(y),$$

[4] On the general analysis with the second class constraints see [31].

$$\left\{\int \mathrm{d}^3x\Lambda^j(x)G_j(x), A^i_a(y)\right\} = \int \mathrm{d}^3x\Lambda^j(x)\left\{\partial_b E^b_j + \varepsilon_{mjn}A^m_b E^{bn}, A^i_a(y)\right\} = \\ = \gamma\partial_a\Lambda^i(y) + \gamma\varepsilon_{mji}\Lambda^j(y)A^m_a(y).$$

Hence,

$$\begin{aligned}
&\left\{G(\Lambda), \partial_a E^a_i(y) + \varepsilon_{jik}A^j_a E^{ak}(y)\right\} = \\
&= \gamma\varepsilon_{ijn}\partial_a\left(\Lambda^j(y)E^{an}(y)\right) - \gamma\varepsilon_{jik}A^j_a(y)\varepsilon_{k\ell n}\Lambda^\ell(y)E^{an}(y) + \\
&+ \gamma\varepsilon_{jik}E^{ak}\left(\partial_a\Lambda^j(y) + \varepsilon_{m\ell j}\Lambda^\ell(y)A^m_a(y)\right) = \\
&= \gamma\varepsilon_{ijn}\Lambda^j(y)\partial_a E^{an}(y) + \gamma\left(\varepsilon_{ijk}\varepsilon_{\ell nk} - \varepsilon_{ink}\varepsilon_{\ell jk}\right)\Lambda^\ell(y)E^{an}(y)A^j_a(y) = \\
&= \gamma\varepsilon_{i\ell k}\Lambda^\ell(y)\partial_a E^{ak}(y) + \gamma\varepsilon_{i\ell k}\Lambda^\ell(y)\varepsilon_{jkn}E^{an}(y)A^j_a(y) = \gamma\varepsilon_{i\ell k}\Lambda^\ell(y)G_k(y).
\end{aligned}$$

Finally, smearing also the second term in the above bracket we get

$$\{G(\Lambda_1), G(\Lambda_2)\} = \frac{\gamma}{2}G\left([\Lambda_1, \Lambda_2]\right), \tag{105}$$

which we recognize as the $su(2)$ algebra structure equations.

We should not be surprised of the appearance of this extra constraint. When we use the tetrad formalism, we introduce a new symmetry in the theory, the invariance under local gauge transformation. The Gauss constraint is there to enforce this invariance at the canonical level. On the other hand, it might be more puzzling that although the local gauge invariance of the covariant action was the full Lorentz group, the Legendre transform (100) is only invariant under SU(2).

The origin of this puzzle lies precisely in the change of variables (98-99). The Ashtekar-Barbero connection is an $SU(2)$ connection, not a Lorentz connection. In particular, it is not the pull-back of the space-time Lorentz connection [32]. It should then be seen as an auxiliary variable, useful to recast the algebra in a first class form.

Summarizing, in this formulation of General Relativity the theory is described by an extended phase space of dimension $18\cdot\infty^3$ with the fundamental Poisson bracket

$$\left\{A^i_a(x), E^b_j(y)\right\} = \gamma\delta^b_a\delta^i_j\delta^3(x,y) \tag{106}$$

This new internal index i corresponds to the adjoint representation of SU(2), and we can recover the old $12\cdot\infty^3$ dimensional phase space on the constraint surface $G_i = 0$ dividing by gauge orbits generated by G. This group SU(2) should be seen as an auxiliary local symmetry group, since the connection with the original Lorentz group of the tetrad formulation is hidden in the change of variables (98-99). It is only for the special case $\gamma = i$ (which is the original one introduced by Ashtekar) that the relation is manifest: in that case, the SU(2) corresponds to the self-dual subgroup of the Lorentz group. Notice that this case also leads to a simplification of the Hamiltonian constraint (103). On the other hand, the variables are now complex, and to recover general relativity one needs to impose reality conditions. These are particularly difficult to deal with at the quantum level, and for this reason most developments in LQG have focused on γ real. This is what we do also in this review.

3.3 Smearing of the algebra

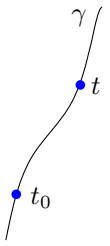
The next step is to smear the algebra (106), as we did previously with the ADM variables. This is needed in order to proceed with the quantization. At this stage, the different tensorial nature of A_a and E^a plays a key role. If we were doing gauge theory on flat spacetime, we would not pay attention

to the different indices, and just smear the connection and the electric field with the same type of test functions. On the contrary, now that the non-trivial geometry of spacetime is our main goal, we have to be careful. In fact, a brief look at (98) shows that the densitised triad is a 2-form. Hence, it is natural to smear it on a surface,

$$E_i(S) \equiv \int_S n_a E_i^a \mathrm{d}^2\sigma, \tag{107}$$

where $n_a = \varepsilon_{abc}\frac{\partial x^b}{\partial \sigma_1}\frac{\partial x^c}{\partial \sigma_2}$ is the normal to the surface. The quantity $E_i(S)$ is the *flux* of E across S.

The connection on the other hand is a 1-form, so it is natural to smear it along a 1-dimensional path. Recall that a connection defines a notion of parallel transport of the fiber over the base manifold. Consider a path γ and a parametrization of it $x^a(s) : [0,1] \to \Sigma$. Given a connection A_a^i we can associate to it an element of $SU(2)$ $A_a \equiv A_a^i \tau_i$ where τ_i are the generator of $SU(2)$.



Then we can integrate A_a along γ as a line integral,

$$A_a^i \longrightarrow \int_\gamma A \equiv \int_0^1 \mathrm{d}s A_a^i(x(s))\frac{\mathrm{d}x^a(s)}{\mathrm{d}s}\tau_i. \tag{108}$$

Next, we define the holonomy of A along γ to be

$$h_\gamma = \mathcal{P}\exp\left(\int_\gamma A\right), \tag{109}$$

where $\mathcal{P}$ stands for the path-ordered product. That is,

$$h_\gamma = \sum_{n=0}^{\infty} \iiint_{1>s_n>\cdots>s_1>0} A(\gamma(s_1))\cdots A(\gamma(s_n))ds_1\cdots ds_n \tag{110}$$

where we parametrized the line with $s \in [0,1]$.

On the definition of the holonomy. More precisely, we call holonomy the solution of the differential equation

$$\frac{d}{dt}h_\gamma(t) - h_\gamma(t)A(\gamma(t)) = 0, \qquad h_\gamma(0) = 1. \tag{111}$$

If we integrate the equation by iteration we have

$$\begin{aligned} h_\gamma(t) &= 1 + \int_0^t A(\gamma(s))h_\gamma(s)ds = \\ &= 1 + \int_0^t A(\gamma(s_1))h_\gamma(s_1)ds_1 + \int_0^t\int_{s_1}^1 A(\gamma(s_1))A(\gamma(s_2))h_\gamma(s_2)ds_1ds_2 = \\ &= \dots \end{aligned} \tag{112}$$

therefore formally

$$h_\gamma(t) = \sum_{n=0}^{\infty} \iiint_{t>s_1>\cdots>s_n>0} A(\gamma(s_1))\cdots A(\gamma(s_n))ds_1\cdots ds_n. \tag{113}$$

To complete the proof, one needs to show that the series is well defined. Indeed, it converges with respect to sup norm

$$||h_\gamma|| \leq \sum_{n=0}^{\infty} \iiint_{t>s_1>\cdots>s_n>0} ||A(\gamma(s_1))\cdots A(\gamma(s_n))||\, ds_1\cdots ds_n \leq \sum_{n=0}^{\infty} ||A||^n \frac{t^n}{n!}. \tag{114}$$

For further reference, let us also notice that the terms of the series can be written as integrals over square domains $(s_1,\dots,s_n) \in [0,t]^n$, instead triangle domains $t > s_1 > \cdots > s_n > 0$. This gives

$$h_\gamma(t) = \mathcal{P}\exp\left[\oint_\gamma A\right] = \sum_{n=0}^{\infty}\frac{1}{n!}\iiint_\square \mathcal{P}\left(A(\gamma(s_1))\cdots A(\gamma(s_n))\right). \tag{115}$$

Let us list some useful properties of the holonomy.

- The holonomy of the composition of two paths is the product of the holonomies of each path,

$$h_{\beta\alpha} = h_\beta h_\alpha. \tag{116}$$

- Under a local gauge transformations $g(x) \in \mathrm{SU}(2)$, the holonomy transforms as

$$h^g_\gamma = g_{s(\gamma)}\, h_\gamma\, g^{-1}_{t(\gamma)}, \tag{117}$$

  where $s(\gamma)$ and $t(\gamma)$ are respectively the source and target points of the line γ.

- Under the action of diffeomorphism, the holonomy transforms as

$$h_\gamma\left(\phi^* A\right) = h_{\phi\circ\gamma}\left(A\right). \tag{118}$$

- The functional derivative with respect to the connection gives

$$\frac{\delta h_\gamma[A]}{\delta A^i_a(x)} = \begin{cases} \frac{1}{2}\dot{x}^a \delta^{(3)}(\gamma(s),x)\,\tau_i h_\gamma & \text{if } x \text{ is the source of } \gamma \\ \frac{1}{2}\dot{x}^a \delta^{(3)}(\gamma(s),x)\, h_\gamma \tau_i & \text{if } x \text{ is the target of } \gamma \\ \dot{x}^a \delta^{(3)}(\gamma(s),x)\, h_\gamma(0,s)\tau_i h_\gamma(s,1) & \text{if } x \text{ is inside } \gamma \end{cases} \tag{119}$$

Basic proofs. If not otherwise specified, we assume for simplicity that the source is $t_0 = 0$. For two composable paths α and β, we define the composition

$$\beta\alpha = \begin{cases} \beta(t), & \text{if } t \in [0,S] \\ \alpha(t-S), & \text{if } t \in [S,T+S] \end{cases} \tag{120}$$

If we split the integrals as $\int_0^{S+T} = \int_0^S + \int_S^{S+T}$, (115) reads

$$\iiint_\square = \sum_{i=0}^n \binom{n}{i} \int_0^S dt_1 \cdots \int_0^S dt_i \int_0^T dt_{i+1} \cdots \int_0^T dt_n. \tag{121}$$

This allows us to split the path ordering as follows,

$$\mathcal{P}\left(A(\beta\alpha(t_1)) \cdots A(\beta\alpha(t_n))\right) = \mathcal{P}\left(A(\beta(t_{i+1})) \cdots A(\beta(t_n))\right) \mathcal{P}\left(A(\alpha(t_1)) \cdots A(\alpha(t_i))\right). \tag{122}$$

Hence, the holonomy along the composite path is

$$\begin{aligned} h_{\beta\alpha} &= \sum_{n=0}^\infty \sum_{i=0}^n \frac{1}{i!\,(n-i)!} \int_{\square(T)} \mathcal{P}\left(A(\beta(t_{i+1})) \cdots A(\beta(t_n))\right) \int_{\square(S)} \mathcal{P}\left(A(\alpha(t_1)) \cdots A(\alpha(t_i))\right) \\ &= \sum_{i,q=0}^\infty \frac{1}{q!} \int_{\square(T)} \mathcal{P}\left(A(\beta(t_1)) \cdots A(\beta(t_q))\right) \frac{(-1)^i}{i!} \int_{\square(S)} \mathcal{P}\left(A(\alpha(t_1)) \cdots A(\alpha(t_i))\right) \\ &= h_\beta h_\alpha, \end{aligned}$$

which proves (116). In the second step, we defined $q = n - i$.

To prove (117), let us introduce families of vectors $u(t)$ and $w(t)$ such that

$$w(t) = u(t) g(\gamma(t)) = h_\gamma(t) g(\gamma(t)) u(0). \tag{123}$$

$w(t)$ satisfies the following differential equation,

$$\begin{aligned}\frac{d}{dt}w(t) &= h\dot{g}u(0) + \dot{h}gu(0) = hgg^{-1}\dot{g}u(0) + hgg^{-1}Agu(0) \\ &= hg\left(g^{-1}\dot{g} + g^{-1}Ag\right)u(0) = \left(g\dot{g}^{-1} + g^{-1}Ag\right)ghu(0) = A^g w(t),\end{aligned}$$

which implies that $w(t) = h^g_\gamma(t)w(0) = h^g_\gamma(t)g(\gamma(0))u(0)$. Comparing with (123) we find (117).

The action of a diffeomorphism ϕ on the line integral of the connection is given by

$$\int_\gamma \phi A = \int_0^1 A_\mu\left(\phi\circ\gamma(t)\right)\partial_\nu\phi^\mu\left(\gamma(t)\right)\dot{\gamma}^\nu(t)dt = \int_0^1 A_\mu\left(\phi\circ\gamma(t)\right)\frac{d}{dt}\left(\phi\circ\gamma\right)^\mu(t)dt, \tag{124}$$

which implies immediately

$$h_\gamma\left(\phi A\right) = h_{\phi\circ\gamma}\left(A\right). \tag{125}$$

Finally, to prove (119), let us compute the differential equation satisfied by $\frac{\delta h_\gamma(t_0,t)}{\delta A^i_a(x)}$ in the two cases of x being inside the path or at one boundary. If x is inside γ, from (111) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta h_\gamma(t_0,t)}{\delta A^i_a(x)} - \frac{\delta h_\gamma(t_0,t)}{\delta A^i_a(x)}A(\gamma(t)) = 0, \tag{126}$$

which is solved by $\frac{\delta h_\gamma(t_0,t)}{\delta A^i_a(x)} = h(t_0,s)\tau_i\dot{x}^a h(s,t)$. Then, by the Leibniz rule we also know that

$$\frac{\delta h_\gamma(t_0,t)}{\delta A^i_a(x)} = \frac{\delta h_\gamma(t_0,s)}{\delta A^i_a(x)}h_\gamma(s,t) + h_\gamma(t_0,s)\frac{\delta h_\gamma(s,t)}{\delta A^i_a(x)} = h(t_0,s)\tau_i\dot{x}^a h(s,t). \tag{127}$$

From this we can argue that if x is at one boundary we must have

$$\frac{\delta h_\gamma(t_0,t)}{\delta A^i_a(x)} = \begin{cases} \frac{1}{2}\dot{x}^a\tau_i h_\gamma(t_0,t) & \text{if } x \text{ is the source of } \gamma \\ \frac{1}{2}\dot{x}^a h_\gamma(t_0,t)\tau_i & \text{if } x \text{ is the target of } \gamma \end{cases} \tag{128}$$

3.4 Summary

In this section we have taken the two key steps needed to prepare general relativity for the loop quantization. The first step was to reformulate the theory in terms of the tetrad field and an SU(2) independent connection, the Ashtekar-Barbero connection (99). The second step was to regularize the resulting Poisson algebra using paths and surfaces, instead of the all of space as in traditional smearings (cf. what was done in the ADM formulation). The resulting smeared algebra of $h_\gamma[A]$ and $E_i(S)$ is called *holonomy-flux algebra*. It provides the most natural regular (i.e. no delta functions appear) version of the Poisson algebra (106).

4 Loop quantum gravity: Kinematics

The formulation of General Relativity in terms of the Ashtekar-Barbero connection and the densitized triad let us to talk about General Relativity in the language of a $SU(2)$ gauge theory, with Poisson brackets (106) and the three sets of constraints

$G_i = 0$	Gauss law
$H_a = 0$	Spatial diffeomorphism invariance
$H = 0$	Hamiltonian constraint

The difference with a gauge theory is of course in the dynamics: In gauge theory, after imposing the Gauss law, we have a physical Hamiltonian. Here instead we still have a fully constrained system.

Let us first briefly review some basic steps in the quantization of gauge theories. In a nutshell, the usual procedure goes along these lines:

- Use the Minkowski metric to define a Gaussian measure δA on the space of connections modulo gauge-transformations.
- Consider the Hilbert space $L_2(A, \delta A) \ni \psi[A]$ and define the Schrödinger representation

$$\hat{A}^i_a \psi[A] = A^i_a \psi[A], \tag{129a}$$

$$\hat{E}^a_i \psi[A] = -i\hbar\gamma \frac{\delta}{\delta A^i_a} \psi[A], \tag{129b}$$

  which satisfies the canonical commutation relation,

$$\left[\hat{A}^i_a(x), \hat{E}^b_j(y)\right] = i\hbar\gamma \delta^b_a \delta^i_j \delta^3(x, y). \tag{130}$$

- Impose the Gauss law constraint, $\hat{G}_i \psi[A] = 0$, which selects the gauge-invariant states.
- Study the dynamics with the physical Hamiltonian,

$$i\partial_t \psi[A] = \hat{H}\psi[A]. \tag{131}$$

  In practice, the easiest way to deal with the dynamics is to decompose A in plane waves and the Hilbert space into a Fock space, $L_2(A, \delta A) = \mathcal{H}_{Fock} = \bigoplus_n \mathcal{H}_n$, the direct sum over n-particles spaces.

The key difference in the case of general relativity is that we do not have a background metric at disposal to define the integration measure, since now the metric is a fully dynamical quantity. Hence, we need to define a measure on the space of connections without relying on any fixed background metric. The key to do this is the notion of cylindrical functions, which we introduce next.

Commutator.

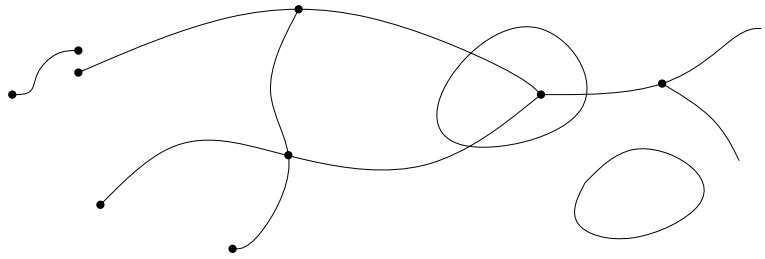
$$\begin{aligned}\left[\hat{A}^i_a(x), \hat{E}^b_j(y)\right] \psi[A] &= -i\hbar\gamma A^i_a(x) \frac{\delta}{\delta A^j_b(y)} \psi[A] + i\hbar\gamma \frac{\delta}{\delta A^j_b(y)} A^i_a(x) \psi[A] \\ &= i\hbar\gamma \frac{\delta A^i_a(x)}{\delta A^j_b(y)} \psi[A] = i\hbar\gamma \delta^b_a \delta^i_j \delta^3(x, y) \psi[A]\end{aligned}$$

4.1 Cylindrical functions and the kinematical Hilbert space

Roughly speaking, a cylindrical function is a functional of a field that depends only on some subset of components of the field itself. In the case at hand, the field is the connection, and the cylindrical functions are functionals that depend on the connection only through the holonomies $h_e[A] = \mathcal{P}\exp\left(\int_e A\right)$ along some finite set of paths e. Consider a graph Γ, defined as a collection of oriented paths $e \subset \Sigma$ (we will call these paths the *links* of the graph) meeting at most at their endpoints. Given a graph $\Gamma \subset \Sigma$ we denote by L the total number of links that it contains. A cylindrical function is a couple (Γ, f) of a graph and a a smooth function $f : SU(2)^L \longrightarrow \mathbb{C}$, and it is given by a functional of the connection defined as

$$\langle A \mid \Gamma, f\rangle = \psi_{(\Gamma,f)}[A] = f(h_{e_1}[A], \ldots, h_{e_L}[A]) \in Cyl_\Gamma \tag{132}$$

where e_i with $i = 1, \ldots, L$ are the links of the corresponding graph Γ.



Collection of paths $\Gamma = \{e_1, \ldots, e_n\}$

This space of functionals can be turned into an Hilbert space if we equip it with a scalar product. The switch from the connection to the holonomy is crucial in this respect, because the holonomy is an element of SU(2), and the integration over SU(2) is well-defined. In particular, there is a unique gauge-invariant and normalized measure $\mathrm{d}h$, called the Haar measure. Using L copies of the Haar measure, we define on Cyl_Γ the following scalar product,

$$\left\langle \psi_{(\Gamma,f)} \mid \psi_{(\Gamma,f')} \right\rangle \equiv \int \prod_e \mathrm{d}h_e \overline{f(h_{e_1}[A], \ldots, h_{e_L}[A])} f'(h_{e_1}[A], \ldots, h_{e_L}[A]). \tag{133}$$

This turns Cyl_Γ into an Hilbert space $\mathcal{H}_\Gamma$ associated to a given graph Γ.

Next, we define the Hilbert space of all cylindrical functions for all graphs as the direct sum of Hilbert spaces on a given graph,

$$\mathcal{H}_{kin} = \bigoplus_{\Gamma \subset \Sigma} \mathcal{H}_\Gamma. \tag{134}$$

The scalar product on $\mathcal{H}_{kin}$ is easily induced from (133) in the following manner: if ψ and ψ' share the same graph, then (133) immediately applies. If they have different graphs, say Γ_1 and Γ_2, we consider a further graph $\Gamma_3 \equiv \Gamma_1 \cup \Gamma_2$, we extend f_1 and f_2 trivially on Γ_3, and define the scalar product as (133) on Γ_3:

$$\left\langle \psi_{(\Gamma_1,f_1)} \mid \psi_{(\Gamma_2,f_2)} \right\rangle \equiv \left\langle \psi_{(\Gamma_1\cup\Gamma_2,f_1)} \mid \psi_{(\Gamma_1\cup\Gamma_2,f_2)} \right\rangle. \tag{135}$$

The key result, due to Ashtekar and Lewandowski, is that (134) defines an Hilbert space over (suitably generalized, see [35] for details) gauge connections A on Σ, i.e.

$$\mathcal{H}_{kin} = L_2[A, \mathrm{d}\mu_{AL}]. \tag{136}$$

The integration measure $\mathrm{d}\mu_{AL}$ over the space of connections is called the Ashtekar-Lewandowski measure. What (136) means is that (135) can be seen as a scalar product between cylindrical functionals of the

connection with respect to the Ashtekar-Lewandowski measure:

$$\int \mathrm{d}\mu_{AL}\, \overline{\psi_{(\Gamma_1,f_1)}(A)}\psi_{(\Gamma_2,f_2)}(A) \equiv \left\langle \psi_{(\Gamma_1,f_1)} \mid \psi_{(\Gamma_2,f_2)} \right\rangle . \tag{137}$$

Now that we have a candidate kinematical Hilbert space which does not require a background metric, let us look for a representation of the holonomy-flux algebra on it. To that end, it is convenient to introduce an orthogonal basis in the space. This can be done easily thanks to the Peter-Weyl theorem. It states that a basis on the Hilbert space $L_2(G, \mathrm{d}\mu_{Haar})$ of functions on a compact group G is given by the matrix elements of the unitary irreducible representation of the group. For the case of $SU(2)$,

$$f(g) = \sum_j \hat{f}^j_{mn} D^{(j)}_{mn}(g) \qquad \begin{matrix} j = 0, \frac{1}{2}, 1, \dots \\ m = -j, \dots, j \end{matrix} \tag{138}$$

where the Wigner matrices $D^{(j)}_{mn}(g)$ give the spin-j irreducible matrix representation of the group element g. This immediately applies to $\mathcal{H}_\Gamma$, since the latter is just a tensor product of $L_2(SU(2), \mathrm{d}\mu_{Haar})$. That is, the basis elements are

$$\langle A \mid \Gamma; j_e, m_e, n_e \rangle \equiv D^{(j_1)}_{m_1 n_1}(h_{e_1}) \dots D^{(j_n)}_{m_n n_n}(h_{e_n}), \tag{139}$$

and a function $\psi_{(\Gamma,f)}[A] \in \mathcal{H}_\Gamma$ can be decomposed as

$$\psi_{(\Gamma,f)}[A] = \sum_{j_e, m_e, n_e} \hat{f}^{j_1,\dots,j_n}_{m_1,\dots,m_n,n_1,\dots,n_n} D^{(j_1)}_{m_1 n_1}(h_{e_1}[A]) \dots D^{(j_n)}_{m_n n_n}(h_{e_n}[A]). \tag{140}$$

On this basis, we can give a Schrödinger representation like (129) for the regularized holonomy-flux version of the algebra. Consider for simplicity the fundamental representation, $h_e \equiv D^{(\frac{1}{2})}(h_e)$. The holonomy acts by multiplication,

$$\hat{h}_\gamma[A] h_e[A] = h_\gamma[A] h_e[A], \tag{141a}$$

and the flux through the derivative (119),

$$\hat{E}_i(S) h_e[A] = -i\hbar\gamma \int_S \mathrm{d}^2\sigma n_a \frac{\delta h_e[A]}{\delta A^i_a(x(\sigma))} = \pm i\hbar\gamma h_{e_1}[A] \tau_i h_{e_2}[A]. \tag{141b}$$

Here e_1 and e_2 are the two new edges defined by the point at which the triad acts and the sign depends on the relative orientation of e and S. The action vanishes, $\hat{E}[S] h_{e_1}[A] = 0$, when e is tangential to S or $e \cap S = 0$.

Action of the Flux. In the case $e \cup S = p$ recalling (119) we have

$$\hat{E}_i(S) h_e = -i\hbar\gamma \int_S \mathrm{d}^2\sigma n_a \int_0^1 \mathrm{d}s \dot{x}^a(s) \delta^3(p, x(\sigma)) h_{e_1} \tau_i h_{e_2} = \tag{142}$$

$$-i\hbar\gamma(-1)^{o(p)} h_e(0,s) \tau_i h_e(s,1) \tag{143}$$

where $(-1)^{o(p)} = \pm 1$ depending if σ_1, σ_2, s is a right or left handed coordinates system

$$\int_S \int_0^1 \mathrm{d}\sigma_1 \mathrm{d}\sigma_2 \mathrm{d}s \varepsilon_{abc} \frac{\partial x^a}{\partial \sigma_1} \frac{\partial x^b}{\partial \sigma_2} \frac{\partial x^c}{\partial s} \delta^3(x(s), x(\sigma)) = \pm \int \mathrm{d}^3 x \delta(x) = \pm 1 \tag{144}$$

If there are no intersection the integration of the $\delta^3(p, x(\sigma))$ is 0. In the case of e tangent to S we can compute the action of E as a limit of a double intersection. Because the two contributes have relative opposite sign the limit is trivially 0.

Consider now the action of the scalar product of two fluxes acting inside the link,

$$\hat{E}_i(S)\hat{E}^i(S)h_e[A] = -\hbar^2\gamma^2\, h_{e_1}[A]\tau^i\tau_i h_{e_2}[A]. \tag{145}$$

On the right hand side, we see the appearance of the scalar contraction of algebra generators, $\tau^i\tau_i \equiv C^2$. This scalar product is known as the Casimir operator of the algebra. In the fundamental representation considered here, $C^2 = -\frac{3}{4}\mathbb{1}_2$. The Casimir clearly commutes with all group elements, thus (145) can be written as

$$\hat{E}_i(S)\hat{E}^i(S)h_e[A] = -\hbar^2 C^2\gamma^2\, h_{e_1}[A]h_{e_2}[A] = -\hbar^2 C^2\gamma h_e[A]. \tag{146}$$

This expression will be useful below.

On the other hand, if two consecutive fluxes act on one endpoint, say the target, we get

$$\hat{E}_i(S)\hat{E}_j(S)h_e[A] = -\hbar^2\gamma^2\, h_e[A]\tau_i\tau_j. \tag{147}$$

From this result we immediately find that two flux operators *do not commute*,

$$[\hat{E}_i(S), \hat{E}_j(S)]h_e[A] = -\hbar^2\gamma^2\, h_e[A][\tau_i, \tau_j] = -\hbar^2\gamma^2\epsilon_{ij}{}^k\, h_e[A]\tau_k. \tag{148}$$

The actions (141) of the holonomy-flux algebra trivially extends to a generic basis element $D^{(j)}(h)$. The action (141a) is unchanged, and in the right hand side of (141b) one simply has to replace τ_i by the generator J_i in the arbitrary irreducible j. Consequently, in (146) we have the Casimir $C_j^2 = -j(j+1)\mathbb{1}_{2j+1}$ on a generic irreducible representation,

$$\hat{E}_i(S)\hat{E}^i(S)D^{(j)}(h_e) = \hbar^2\gamma^2 j(j+1)\, D^{(j)}(h_e). \tag{149}$$

Finally, the action is extended by linearity over the whole $\mathcal{H}_{kin}$. The remarkable fact is that this representation of the holonomy-flux algebra on $\mathcal{H}_{kin}$ is *unique*, as proved by Fleischhack and Lewandowski, Okolow, Sahlmann, Thiemann [33]. This uniqueness result can be compared to the Von Neumann theorem in quantum mechanics on the uniqueness of the Schrödinger representation. It is well-known that the uniqueness does not extend to interacting field theories on flat spacetime. Remarkably, insisting on background-independence reintroduces such uniqueness also for a field theory.

What we have accomplished with this construction is the definition of a well-behaved kinematical Hilbert space for general relativity. It carries a representation of the canonical Poisson algebra, and as a bonus, this representation is unique. Following Dirac, we now have a well-posed problem of reduction by the constraints:

$$\mathcal{H}_{kin} \quad \xrightarrow{\hat{G}_i = 0} \quad \mathcal{H}^0_{kin} \quad \xrightarrow{\hat{H}^a = 0} \quad \mathcal{H}_{Diff} \quad \xrightarrow{\hat{H} = 0} \quad \mathcal{H}_{phys}. \tag{150}$$

4.2 Gauge-invariant Hilbert space

The first step is to find the solutions of the quantum Gauss constraint. These are the states in $\mathcal{H}_{kin}$ that are $SU(2)$ gauge invariant. These solutions define a new Hilbert space, that we call $\mathcal{H}^0_{kin}$ where we leave the subindex *kin* to keep in mind that there are still constraints to be solved before arriving to $\mathcal{H}_{phys}$. The action of the Gauss constraint is easily represented in $\mathcal{H}_{kin}$. In fact, recall that under gauge transformations

$$h_e \longrightarrow h'_e = \hat{U}_G h_e = g_{s(e)} h_e g^{-1}_{t(e)}. \tag{151}$$

Similarly, in a generic irrep j we have

$$D^{(j)}(h_e) \longrightarrow D^{(j)}(h'_e) = D^{(j)}(g_{s(e)} h_e g_{t(e)}^{-1}) = D^{(j)}(g_{s(e)}) D^{(j)}(h_e) D^{(j)}(g_{t(e)}^{-1}). \tag{152}$$

From this it follows that gauge transformations act on the source and targets of the links, namely on the *nodes* of a graph. Imposing gauge-invariance then means requiring the cylindrical function to be invariant under action of the group at the nodes:

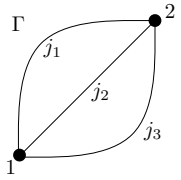
$$f_0(h_1, \dots, h_L) \equiv f_0(g_{s_1} h_1 {g_{t_1}}^{-1}, \dots, g_{s_L} h_L {g_{t_L}}^{-1}) \tag{153}$$

This property can be easily implemented via *group averaging*: given an arbitrary $f \in Cyl_\Gamma$, the function

$$f_0(h_1, \dots, h_L) \equiv \int \prod_n \mathrm{d}g_n \; f(g_{s_1} h_1 {g_{t_1}}^{-1}, \dots, g_{s_L} h_L {g_{t_L}}^{-1}) \tag{154}$$

clearly satisfies (153).

Example: The theta graph. To give a constructive example, let us consider the following graph,



A generic cylindrical function can be expressed in terms of the orthonormal basis using (140). Since the gauge transformations act only on the group elements, the gauge-invariant part is obtained looking at the gauge-invariant part of the product of Wigner matrices,

$$f_{inv}(h_1, \dots, h_3) = \sum_{j_l, m_l, n_l} \hat{f}^{j_1, j_2, j_3}_{m_1, m_2, m_3, n_1, n_2, n_3} \left[D^{(j_1)}_{m_1 n_1}(h_1) D^{(j_2)}_{m_2 n_2}(h_2) D^{(j_3)}_{m_3 n_3}(h_3) \right]_{inv}.$$

Using the definition (154), the invariant part of the basis is

$$\begin{aligned} &\left[D^{(j_1)}_{m_1 n_1}(h_1) D^{(j_2)}_{m_2 n_2}(h_2) D^{(j_3)}_{m_3 n_3}(h_3) \right]_{inv} = \\ &= \int \mathrm{d}g_1 \mathrm{d}g_2 D^{(j_1)}_{m_1 n_1}(g_1 h_1 g_2^{-1}) D^{(j_2)}_{m_2 n_2}(g_1 h_2 g_2^{-1}) D^{(j_3)}_{m_3 n_3}(g_1 h_3 g_2^{-1}) = \\ &= \mathcal{P}_{m_1 m_2 m_3 \alpha_1 \alpha_2 \alpha_3} \mathcal{P}_{\beta_1 \beta_2 \beta_3 n_1 n_2 n_3} D^{(j_1)}_{\alpha_1 \beta_1}(h_1) D^{(j_2)}_{\alpha_2 \beta_2}(h_2) D^{(j_3)}_{\alpha_3 \beta_3}(h_3), \end{aligned}$$

where $\mathcal{P}_{m_1 m_2 m_3 \alpha_1 \alpha_2 \alpha_3}$ is the projector on the gauge invariant space,

$$\mathcal{P}_{m_1 m_2 m_3 \alpha_1 \alpha_2 \alpha_3} = \int \mathrm{d}g_1 D^{(j_1)}_{m_1 \alpha_1}(g_1) D^{(j_2)}_{m_2 \alpha_2}(g_1) D^{(j_3)}_{m_3 \alpha_3}(g_1).$$

This projector can be written in terms of normalized Clebsch-Gordan coefficients, or Wigner's 3j-m symbols, as

$$\int \mathrm{d}g_1 D^{(j_1)}_{m_1 \alpha_1}(g_1) D^{(j_2)}_{m_2 \alpha_2}(g_1) D^{(j_3)}_{m_3 \alpha_3}(g_1) = \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \overline{\begin{pmatrix} j_1 & j_2 & j_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix}}. \tag{155}$$

With this notation,

$$\begin{aligned}
&\left[D^{(j_1)}_{m_1n_1}(h_1)D^{(j_2)}_{m_2n_2}(h_2)D^{(j_3)}_{m_3n_3}(h_3)\right]_{inv} = \\
&= \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \overline{\begin{pmatrix} j_1 & j_2 & j_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix}} \begin{pmatrix} j_1 & j_2 & j_3 \\ \beta_1 & \beta_2 & \beta_3 \end{pmatrix} \overline{\begin{pmatrix} j_1 & j_2 & j_3 \\ n_1 & n_2 & n_3 \end{pmatrix}} D^{(j_1)}_{\alpha_1\beta_1}(h_1)D^{(j_2)}_{\alpha_2\beta_2}(h_2)D^{(j_3)}_{\alpha_3\beta_3}(h_3) \\
&= \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \overline{\begin{pmatrix} j_1 & j_2 & j_3 \\ n_1 & n_2 & n_3 \end{pmatrix}} \prod_e D^{(j_e)}(h_e) \prod_n i_n
\end{aligned}$$

where i_n is a short-hand notation for the 3j-m symbols. Notice that these are the invariant tensors in the space of $\bigotimes_{e\in n} j_e$ of all the spins that enters in the node n. Finally, we have

$$\begin{aligned}
f_{inv} &= \sum_{j_e} \prod_e D^{(j_e)}(h_e) \prod_n i_n \sum_{m_e n_e} \hat{f}^{j_1,j_2,j_3}_{m_1,m_2,m_3,n_1,n_2,n_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \overline{\begin{pmatrix} j_1 & j_2 & j_3 \\ n_1 & n_2 & n_3 \end{pmatrix}} \\
&= \sum_{j_e} \hat{f}^{j_1,j_2,j_3} \prod_e D^{(j_e)}(h_e) \prod_n i_n,
\end{aligned} \tag{156}$$

with the new coefficients $\hat{f}^{j_1,j_2,j_3}$ including the sums over the magnetic numbers m_e, n_e.

The group averaging amounts to inserting on each node n the following projector,

$$\mathcal{P} = \int \mathrm{d}g \prod_{e\in n} D^{(j_e)}(g). \tag{157}$$

Here the integrand is an element in the tensor product of $SU(2)$ irreducible representations,

$$\prod_e D^{(j_e)}_{m_e n_e}(h_e) \in \bigotimes_e V^{(j_e)}. \tag{158}$$

As such, it transforms non-trivially under gauge transformation and is in general reducible,

$$\bigotimes_e V^{(j_e)} = \bigoplus_i V^{(j_i)}. \tag{159}$$

Then, the integration in (157) selects the gauge invariant part of $\bigotimes_e V^{(j_e)}$, namely the *singlet* space $V^{(0)}$, if the latter exists. Since $\mathcal{P}$ is a projector, we can decompose it in terms of a basis of $V^{(0)}$. Denoting i_α a vector ("ket") in this basis, $\alpha = 1, \ldots, \mathrm{dim}V^{(0)}$, and i^*_α the dual ("bra"),

$$\mathcal{P} = \sum_{\alpha=1}^{\mathrm{dim}V^{(0)}} i_\alpha \, i^*_\alpha. \tag{160}$$

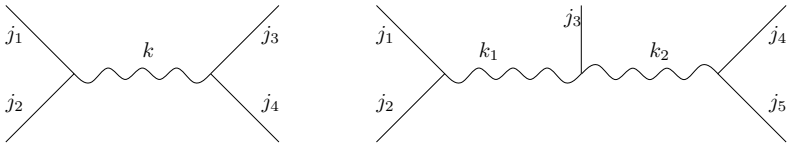
These invariants are called *intertwiners*. For the case of a 3-valent node as in the above example, $\mathrm{dim}V^{(0)} = 1$ and the unique intertwiner i is given by Wigner's 3j-m symbols (cf. (155)). More precisely in the case of a three-valent node the space

$$\left[V^{(j_1)} \otimes V^{(j_2)} \otimes V^{(j_3)}\right]_{inv}$$

is non-empty only when the following Clebsch-Gordan conditions hold,

$$|j_2 - j_3| \leq j_1 \leq j_2 + j_3. \tag{161}$$

For an n-valent node, the space $V^{(0)}$ can have a larger dimension. To visualize the intertwiners, it is convenient to add first two irreps only, then the third, and so on. This gives rise to a decomposition over virtual links, which for $n = 4$ and $n = 5$ looks as follows:



The virtual spins k_i label the intertwiners.

Brief proof that $\mathcal{P}$ is a projector. First we check that $\mathcal{P}\mathcal{P} = \mathcal{P}$

$$\mathcal{P}\mathcal{P} = \int \mathrm{d}g_1 \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_1) \prod_{e\in n} D^{(j_e)}(g_2) = \int \mathrm{d}g_1 \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_1) D^{(j_e)}(g_2) = \tag{162}$$

$$= \int \mathrm{d}g_1 \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_1 g_2) = \int \mathrm{d}g_2 \int \mathrm{d}g_1 \prod_{e\in n} D^{(j_e)}(g_1) = \mathcal{P} \tag{163}$$

than that is left and right invariant

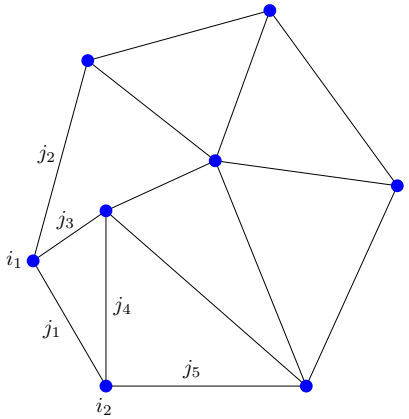
$$\prod_{e\in n} D^{(j_e)}(g_1)\mathcal{P} = \int \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_1) D^{(j_e)}(g_2) = \int \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_1 g_2) = \mathcal{P} \tag{164}$$

$$\mathcal{P} \prod_{e\in n} D^{(j_e)}(g_1) = \int \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_2) D^{(j_e)}(g_1) = \int \mathrm{d}g_2 \prod_{e\in n} D^{(j_e)}(g_2 g_1) = \mathcal{P} \tag{165}$$

The facts that $\mathcal{P}$ acts only on the nodes of the graph that label the basis of $\mathcal{H}_{kin}$ and equation (160) implies that the result of the action of $\mathcal{P}$ on elements of $\mathcal{H}_{kin}$ can be written as a linear combination of products of representation matrices $D^{(j)}(h_e)$ contracted with intertwiners, generalizing the result (156). The states labeled with a graph Γ, with an irreducible representation $D^{(j)}(h)$ of spin-j of the holonomy h along each link, and with an element i of the intertwiner space $\mathcal{H}_n \equiv \mathrm{Inv}[\underset{e\in n}{\otimes} V^{(j_e)}]$ in each node, are called spin network states, and are given by

$$\psi_{(\Gamma, j_e, i_n)}[h_e] = \underset{e}{\otimes} D^{(j_e)}(h_e) \underset{n}{\otimes} i_n. \tag{166}$$

Here the indices of the matrices and of the interwiners are hidden for simplicity of notation. Their contraction pattern can be easily reconstructed from the connectivity of the graph.



To complete the discussion, let us show that imposing the gauge-invariance amounts to solving the

Gauss law constraint, $\hat{G}_i\psi = 0$. Consider a gauge invariant node n, and a surface S centered in n of radius ϵ. The action of the total flux operator through S on n vanishes identically:

$$\lim_{\epsilon\to 0} \hat{E}(S)\,|n\rangle = 0. \tag{167}$$

In fact, using (141b) at each link one notices that (167) produces the infinitesimal gauge transformation $g_\alpha = 1 - \alpha_i\tau^i \in SU(2)$ at the node, and because the node is gauge invariant such action vanishes.

Summarizing, spin network states (166) form a complete basis of the Hilbert space of solutions of the quantum Gauss law, $\mathcal{H}^0_{kin}$. The structure of this space is nicely organized by the spin networks basis. As before, different graphs Γ select different orthogonal subspaces, thus $\mathcal{H}^0_{kin}$ decomposes as a direct sum over spaces on a fixed graph,

$$\mathcal{H}^0_{kin} = \bigoplus_{\Gamma\subset\Sigma} \mathcal{H}^0_\Gamma. \tag{168}$$

Furthermore, the Hilbert space on a fixed graph decomposes as a sum over intertwiner spaces,

$$\mathcal{H}^0_\Gamma = L_2[SU(2)^L/SU(2)^N, \mathrm{d}\mu_{Haar}] = \oplus_{j_l}\left(\otimes_n \mathcal{H}_n\right). \tag{169}$$

These two equations are the analogue in loop gravity of the Fock decomposition of the Hilbert space of a free field in Minkowski spacetime into a direct sum of n-particle states, and play an equally important fundamental role.

4.3 Geometric operators

4.3.1 The area operator

The simplest geometric operator that can be constructed in loop quantum gravity is the area operator. The area of a surface S can be given in terms of its normal n_a and the densitized triad E^a_i,

$$A(S) = \int_S \mathrm{d}\sigma_1\mathrm{d}\sigma_2 \sqrt{E^a_i E^{bi} n_a n_b} \tag{170}$$

The area at classical level. We start from the standard definition of area in terms of the metric,

$$A(S) = \int_S \mathrm{d}\sigma_1\mathrm{d}\sigma_2 \sqrt{\det\left(g_{ab}\frac{\partial x^a}{\partial\sigma^\alpha}\frac{\partial x^b}{\partial\sigma^\beta}\right)} \qquad \alpha,\beta = 1,2 \tag{171}$$

using the notation $\partial_1 x^a = \frac{\partial x^a}{\partial\sigma^1}$

$$\det\left(g_{ab}\frac{\partial x^a}{\partial\sigma^\alpha}\frac{\partial x^b}{\partial\sigma^\beta}\right) = g_{ab}g_{cd}\left[\partial_1 x^a\partial_1 x^b\partial_2 x^c\partial_2 x^d - \partial_1 x^a\partial_2 x^b\partial_1 x^c\partial_2 x^d\right] = \tag{172}$$

$$= g_{ab}g_{cd}2\partial_1 x^a\partial_1 x^{[b}\partial_2 x^{c]}\partial_2 x^d = 2g_{a[b}g_{c]d}\partial_1 x^a\partial_1 x^b\partial_2 x^c\partial_2 x^d = \tag{173}$$

$$= g g^{ef} n_e n_f \tag{174}$$

where we recognized that

$$g_{a[b}g_{c]d} = \frac{1}{2}\varepsilon_{ace}\varepsilon_{bdf} g g^{ef} \qquad n_e = \varepsilon_{eab}\frac{\partial x^a}{\partial\sigma^1}\frac{\partial x^b}{\partial\sigma^2} \tag{175}$$

Using the definition of tetrad and densitized tetrad we have

$$A(S) = \int_S \mathrm{d}\sigma_1\mathrm{d}\sigma_2 \sqrt{e^2 e^e_i e^{fi} n_e n_f} = \int_S \mathrm{d}\sigma_1\mathrm{d}\sigma_2 \sqrt{E^e_i E^{fi} n_e n_f} \tag{176}$$

At the quantum level, we know from (141b) that the triad operators act as functional derivatives. The action of the scalar product of two triads operator was also studied, see equation (149), for the case of a surface intersected only once by the holonomy path. The case of a generic graph can be easily dealt with if we regularize the expression for the area in the following way. We introduce a decomposition of S in N two-dimensional cells, and write the integral as the limit of a Riemann sum,

$$A(S) = \lim_{N\to\infty} A_N(S), \tag{177}$$

where the Riemann sum can be expressed as

$$A_N(S) = \sum_{I=1}^{N} \sqrt{E_i(S_I)E^i(S_I)}. \tag{178}$$

Here N is the number of cells, and $E_i(S_I)$ is the flux of E_i through the I-th cell.

Checking the limit. In the limit of infinitesimal cells we have that

$$E_i(S_I) \equiv \int_{S_I} n_a E^a_i \mathrm{d}^2\sigma = \int_{S_I} E^a_i n_a \approx E^a_i n_a S_I \tag{179}$$

In that limit the definition of the area

$$A_N(S) = \sum_{I=1}^{N} \sqrt{E_i(S_I)E^i(S_I)} \approx \sum_{I=1}^{N} \sqrt{E^a_i n^I_a S_I E^{bi} n^I_b S_I} = \sum_{I=1}^{N} S_I \sqrt{E^a_i n^I_a E^{bi} n^I_b} = \tag{180}$$

$$= \int_S \sqrt{E^a_i E^{bi} n_a n_b} \tag{181}$$

Accordingly, we define the area operator as

$$\hat{A}(S) = \lim_{N\to\infty} \hat{A}_N(S), \tag{182}$$

where in $A_N(S)$ we simply replace the classical flux $E_i(S_I)$ by the operator $\hat{E}_i(S_I)$. This operator now acts on a generic spin network state ψ_Γ, where the graph Γ is generic and can intersect S many times. We already know that $\hat{E}_i(S_I)\hat{E}^i(S_I)$ gives zero if S_I is not intersected by any link of the graph. Therefore once the decomposition is sufficiently fine so that each surface S_I is punctured once and only once, taking a further refinement has no consequences. Therefore, the limit amounts to simply sum the contributions of the finite number of punctures p of S caused by the links of Γ. That is,

$$\hat{A}(S)\,\psi_\Gamma = \lim_{N\to\infty} \sum_{I=1}^{N} \sqrt{\hat{E}_i(S_I)\hat{E}^i(S_I)}\,\psi_\Gamma = \sum_{p\in S\cup\Gamma} \hbar\sqrt{\gamma^2 j_p(j_p+1)}\,\psi_\Gamma. \tag{183}$$

There are two key remarks to make to this formula: first of all, the spectrum of the area operator is completely known[5] and *quantized*: the area can only take up discrete values, with minimal excitation being proportional to the squared Planck length $\ell_P^2 = \hbar G$, restoring Newton's constant. This result can be compared with other celebrated quantizations, such as the radii of electron's orbitals in atoms.

Second, the operator has a *diagonal* action on spin networks. Therefore, spin network states are eigenstates of the area operator.

[5] In this expression, we assumed that each puncture is caused by a link crossing the surface. However, it could also happen that S is puncture by a node of the graph. A closed expression for the area spectrum is known also in this general case. See the literature for details [22, 24, 37].

4.3.2 The volume operator

Given a region $R \subset \Sigma$ classically we can define its volume as

$$V(R) = \int_R \mathrm{d}^3x\sqrt{g} = \int_R \mathrm{d}^3x \sqrt{\left|\frac{1}{3!}\varepsilon_{abc}\varepsilon^{ijk}E^a_i E^b_j E^c_k\right|}, \tag{184}$$

where the quantity in absolute value can be recognized as the determinant of the densitised triad, $\det E^a_i$.

Two distinct mathematically well-defined volume operators have been proposed in the literature. One is due to Rovelli and Smolin, and the other to Ashtekar and Lewandowski. We will refer to them as RS and AL respectively. Both of them act non-trivially only at the nodes of a spin network state. Let us begin reviewing the construction by RS.

As we did for the area, we replace the integral over R by the limit of a Riemann sum. Specifically, we consider a partition of the region in cubic cells C_I so that $R \subset \cup_I C_I$, and the integral $\int_R \mathrm{d}^3x$ can be approximated from above by the sum $\sum_I \mathrm{volume}(C_I)$. This partition allows us to rewrite (184) in terms of fluxes. In fact, consider the following integral,

$$W_I \equiv \frac{1}{48}\int_{\partial C_I}\mathrm{d}^2\sigma_1 \int_{\partial C_I}\mathrm{d}^2\sigma_2 \int_{\partial C_I}\mathrm{d}^2\sigma_3 \left|\varepsilon_{ijk}E^a_i(\sigma_1)n_a(\sigma_1)E^b_j(\sigma_2)n_b(\sigma_2)E^c_k(\sigma_3)n_c(\sigma_3)\right|.$$

In the continuum limit where we send the size of the cell $\epsilon \mapsto 0$ and we shrink the cell to a point x, we obtain

$$W_I = \frac{1}{48}\epsilon^{abc}n_a n_b n_c \det E^a_i(x)\,\epsilon^6 \simeq \det E^a_i(x)\,\epsilon^6 \simeq \mathrm{volume}^2(C_I).$$

Hence, we have

$$V(R) = \lim_{\epsilon\mapsto 0}\sum_I \sqrt{W_I}. \tag{185}$$

For the sake of notation, let us subdivide each ∂C_I into surfaces S^α such that $\partial C_I = \cup_\alpha S^\alpha_I$. Then, we can write W_I as a sum of fluxes over three surfaces, and

$$V(R) = \lim_{\epsilon\mapsto 0}\sum_I \sqrt{\frac{1}{48}\sum_{\alpha,\beta,\gamma}\left|\varepsilon_{ijk}E_i(S^\alpha_I)E_j(S^\beta_I)E_k(S^\gamma_I).\right|} \tag{186}$$

Finally, we can simply turn the classical fluxes to operators,

$$\hat{V}(R) = \lim_{\epsilon\mapsto 0}\sum_I \sqrt{\frac{1}{48}\sum_{\alpha,\beta,\gamma}\left|\varepsilon_{ijk}\hat{E}_i(S^\alpha_I)\hat{E}_j(S^\beta_I)\hat{E}_k(S^\gamma_I)\right|}. \tag{187}$$

This is the Rovelli-Smolin volume operator.

As for the area operator, one finds that there exists an "optimal" subdivision, after which the result stays unchanged with any further refinement, and so the limit can be safely taken. For the area, this consisted in the small surfaces being punctured only once at most. Something similar happens for the volume. The "optimal" partition is reached as follows. The nodes of Γ can fall only in the interior of cells, and a cell C_I contains at most one node. In case the cell contains no node, then we assume that it contains at most one link. Moreover, we assume that the partition of the surfaces ∂C_I in cells S^α_I is refined so that links of Γ can intersect a cell S^α_I only in its interior and each cell S^α_I is punctured at most by one link.

This said, let us now study the action of the operator. The first thing to notice is that the presence of the epsilon tensor requires all three fluxes to be different: if two are the same, then their antisymmetric

combination introduced by the epsilon vanishes. In particular, this means that the volume does not act on links, since if no node is present, two of the S_I^α have to be the same. We thus obtain the important result that *the volume operator acts only on nodes of the graph.*

Let us now focus on a single node, i.e. the I-th contribution to (187). We consider the cubic operator

$$\hat{U} = \frac{1}{48} \sum_{\alpha,\beta,\gamma} \left| \varepsilon_{ijk} \hat{E}_i(S^\alpha) \hat{E}_j(S^\beta) \hat{E}_k(S^\gamma) \right|. \tag{188}$$

Let us restrict to gauge-invariant spin networks, where the Gauss law holds and each node is labeled by an intertwiner. First of all, it is immediate to see that the action of (188) on a 3-valent node is zero. In fact, the Gauss law tells us that the sum of the fluxes through a surface around a gauge invariant node is zero. In the 3-valent case, only three S^α give non-zero fluxes, thus

$$\left(\hat{E}_i(S^\alpha) + \hat{E}_i(S^\beta) + \hat{E}_i(S^\gamma) \right) |i\rangle = 0, \tag{189}$$

which implies

$$\hat{E}_i(S^\alpha) |i\rangle = - \left(\hat{E}_i(S^\beta) + \hat{E}_i(S^\gamma) \right) |i\rangle . \tag{190}$$

Using this result in (188) we get zero because two identical fluxes always appear,

$$\varepsilon_{ijk} \hat{E}_i(S^\alpha) \hat{E}_j(S^\beta) \hat{E}_k(S^\gamma) |i\rangle = -\varepsilon_{ijk} \left(\hat{E}_i(S^\beta) + \hat{E}_i(S^\gamma) \right) \hat{E}_j(S^\beta) \hat{E}_k(S^\gamma) |i\rangle = 0.$$

Non-trivial contributions to the volume comes from nodes of valency 4 or higher. Notice that these are the cases for which the intertwiner is not unique, but a genuine independent quantum number. Thus the operator (188) probes precisely the degrees of freedom hidden in the intertwiners. Consider the 4-valent case. Of all the surface cells S^α, only four are punctured by the links. Hence, we have only four contributions to the Gauss law, which we can then use to eliminate one flux in favor of the remaining three,

$$\hat{E}_i(S^4) = -\hat{E}_i(S^1) - \hat{E}_i(S^2) - \hat{E}_i(S^3). \tag{191}$$

The sum over the S^α in (188) then reduces to the sole contributions from the four punctured surfaces. Using (191) these contributions are all equal. A simple combinatorial shows that there are 48 terms, all equal, thus

$$\hat{U} = \left| \varepsilon_{ijk} \hat{E}_i(S^1) \hat{E}_j(S^2) \hat{E}_k(S^3) \right| = \hbar^3 \left| \gamma^3 \varepsilon_{ijk} J_i^1 J_j^2 J_k^3 \right| . \tag{192}$$

In the last step we have used the result (141b), which gives the action of the fluxes in terms of the generators $\vec{J}^a$ in the spin j_a representation. Notice that the orientation factors $\pm$ (between the edge and the surface) are irrelevant because of the modulus.

This is a well-defined cubic operator, whose action spectrum is again discrete, with minimal excitation proportional to $(\hbar G)^{3/2}$. The explicit formula is much more complicated than that of the area. We refer the interested reader to the literature [39].

The alternative proposal by Ashtekar and Lewandowski has the extra property of being sensitive to the differential structure of the graph at the node. The subdivision of the region R into cubic cells is the same as above, the key difference is in how to choose the surfaces to smear the triad filed. Instead of subdividing the boundary ∂C_I into cells S^α, we now consider for each C_I only three surfaces. We assign a local coordinate system x^a, and take the three surfaces to be any three surfaces S_I^a inside the

cube and orthogonal to each other. The Ashtekar and Lewandowski volume operator is defined as

$$\hat{V}(R) = \lim_{\epsilon \mapsto 0} \sum_{C_I} \sqrt{\left| \frac{1}{3!} \varepsilon_{ijk} \varepsilon_{abc} \hat{E}^i(S_I^a) \hat{E}^j(S_I^b) \hat{E}^k(S_I^c) \right|}. \tag{193}$$

Unlike (184), this expression is already written in terms of fluxes, so we do not need the manipulation in terms of W_I. We can directly use (141b) to write

$$\frac{1}{3!} \varepsilon_{ijk} \varepsilon_{abc} \hat{E}^i(S_I^a) \hat{E}^j(S_I^b) \hat{E}^k(S_I^c) = \frac{\hbar^3}{48} \varepsilon_{ijk} \varepsilon_{abc}\, \kappa_a \kappa_b \kappa_c\, J_a^i J_b^j J_c^k, \tag{194}$$

where again $J^i_{e_a}$ is the $SU(2)$ generator in the spin j_a representation, and κ_S is the relative orientation between the cell S_I^a and the edge puncturing it, i.e.

$$\kappa_S(e) = \begin{cases} 1 & \text{if } e \text{ lies above } S \\ -1 & \text{if } e \text{ lies below } S \\ 0 & \text{otherwise} \end{cases} \tag{195}$$

The operator (194) strongly depends on the choice of the triple of surfaces S_I^a. In order to define a volume operator independent from the regulator, we average (194) on all the possible choices of orthogonal triples S_I^a (see the [37] for the details). The result is a sum over the nodes n of the graph,

$$\hat{V}(R) = \hbar^{3/2} \sum_{n \in \Gamma} \sqrt{\left| \frac{\kappa_0}{48} \varepsilon_{ijk} \sum_{e,e',e''} \epsilon(e,e',e'')\, J^i(e) J^j(e') J^k(e'') \right|}, \tag{196}$$

where e,e',e'' runs over the set of edges passing through the node n and ϵ is the orientation function, which equals 0 if the tangent directions of the three edges are linearly dependent and ± 1 if they are linearly independent and oriented positively or negative respect to the orientation of Σ. κ_0 is an undetermined constant introduced by the averaging procedure.

This alternative volume differs from (187) in two ways: firstly, it depends on an arbitrary constant, κ_0, unlike (187) which is unique. Secondly, the absolute value is outside the internal summation, unlike (187) where it is inside. This second differences is what makes the AL operator sensible to the differential structure of the graph: in fact, the action of (196) vanishes on nodes whose edges lie on a plane, because in that case $\epsilon \equiv 0$.

On the other hand, both operators annihilate a 3-valent node, and coincide (up to a proportionality constant) in the 4-valent case, where they are both proportional to $\varepsilon_{ijk} J_i^1 J_j^2 J_k^3$.

Summarizing, both volume operators act only on nodes of the graph. Their matrix elements vanish between different intertwiner spaces, and since every intertwiner space is finite dimensional, their spectra are discrete with minimal excitations proportional to the Planck length cube ℓ_P^3.

Together with the discreteness of the area operator, these results show that in loop quantum gravity the *space geometry is discrete at the Plack scale.* Each spin network describes a *quantum geometry*, where each face dual to a link has an area proportional to the spin j_e, and each region around a node has a volume determined by the intertwiner i_n as well as the spins of the link sharing the node.

It is important to stress that this is not a built-in discretization, as in lattice approaches to quantum gravity. It is a result of the quantum theory, similar to the quantization of the energy levels of an harmonic oscillator, or the radii of the atomic orbitals.

Thanks to this fundamental discreteness, where the minimal geometric excitations are proportional to

the Planck length, the theory is expected not to have ultraviolet divergences, and to resolve the problem of the classical singularities of general relativity at the big bang (e.g. [41]), or at the center of black holes (e.g. [42]). Addressing these issues of course requires taking into account the dynamics of the theory, to which we turn next.

5 Loop quantum gravity: Dynamics

5.1 Solutions of the diffeomorphisms constraint

Spin network states $\psi_{(\Gamma,j_e,i_n)}[A]$ are in $\mathcal{H}^0_{kin}$, i.e. $\hat{G}^i\psi_{(\Gamma,j_e,i_n)}[A] = 0$. The next step in the Dirac program (150) is to implement the spatial diffeomorphisms, namely to find gauge-invariant states such that $\hat{H}^a\psi[A] = 0$.

To that end, consider a finite diffeomorphism ϕ. Its action on the holonomy, as in (118), naturally induces an operator $\hat{\phi}$ on the space of cylindrical functions. This operator maps Cyl_Γ to $Cyl_{\phi\circ\Gamma}$, that is $\hat{\phi}\psi_\Gamma = \psi_{\phi\circ\Gamma}$. Its action is well-defined and unitary, thanks to the fact that the Ashtekar-Lewandowski measure is diffeomorphism invariant. On the other hand, Cyl_Γ and $Cyl_{\phi\circ\Gamma}$ are orthogonal Hilbert spaces, regardless of what the diffeomorphism is. This means that we can not define the action of an "infinitesimal" diffeomorphism diffeomorphisms are all finite, from the perspective of cylindrical functions. Although this might appear as a restriction, it is not a real problem for the construction of $\mathcal{H}_{Diff}$. We can proceed by group averaging as we did for the Gauss constraint, and construct $\mathcal{H}_{Diff}$ from those states invariant under *finite* diffeomorphisms.

However, there are a couple of subtle issues to take into account. The first one has to do with the existence of symmetries of the graphs. Namely, for each graph there are always some diffeomorphisms that act trivially on it, leaving it unchanged. Let us distinguish two cases: the diffeomorphisms that exchange the links among themselves without changing Γ, call them GS_Γ following [40], and those that also preserve each link, and merely shuffle the points inside the link, call them TDiff_Γ. The latter have to be taken out, because their infinite-dimensional trivial action would spoil the group averaging procedure.

The second issue is that unlike imposing the Gauss law, imposing the invariance under diffeomorphism will not result in a subspace of $\mathcal{H}^0_{kin}$, since diffeomorphisms are a non-compact group. Think for instance of $\psi(x) \in L_2[\mathbb{R}, \mathrm{d}x]$ required to be invariant under translations (a non-compact group): the result is a constant function c, which is not in $L_2[\mathbb{R}, \mathrm{d}x]$ since it is not integrable. It defines however a linear functional $c : \psi \in L_2[\mathbb{R}, \mathrm{d}x] \mapsto \mathbb{C}$, since $\int \mathrm{d}x\, c\,\psi(x) = c\,\tilde{\psi}(0)$, the Fourier transform of ψ evaluated in zero. Similarly, solutions to the diffeomorphism constraint can be described in terms of *linear functionals* on $\mathcal{H}^0_{kin}$. Let us denote $\mathcal{H}^{0*}_{kin}$ the space of linear functionals on $\mathcal{H}^0_{kin}$. Then, $\eta \in \mathcal{H}^{0*}_{kin}$ is a diffeomorphic-invariant functional if

$$\eta(\hat{\phi}\psi) = \eta(\psi) \quad \forall\psi \in \mathcal{H}^0_{kin}. \tag{197}$$

The space of such functionals is denoted $\mathcal{H}^*_{Diff}$, and the Hilbert space $\mathcal{H}_{Diff}$ solution of the constraint is constructed by duality. The condition (197) should remind the reader of the analogue condition of gauge-invariance studied in the previous section. There, we were able to implement it via a simple group averaging procedure. In the same manner, we would like to define a projector $\mathcal{P}_{Diff}$ on $\mathcal{H}_{Diff}$ such that

$$\langle\psi|\psi'\rangle_{Diff} \equiv \langle\psi|\mathcal{P}_{Diff}|\psi'\rangle = \sum_{\phi\in \mathrm{Diff}/\mathrm{TDiff}_\Gamma} \langle\hat{\phi}\psi|\psi'\rangle, \tag{198}$$

where the sum is over all the diffeomorphism mapping Γ into Γ' *except* those corresponding to the trivial ones TDiff_Γ. See citeAshtekarReview for more details.

The result of this procedure are spin network states defined on *equivalence classes of graphs under diffeomorphisms*. These equivalence classes are called *knots*, see Figure 1. The study of knots forms an elegant branch of mathematics. The diff-invariant Hilbert space of loop quantum gravity is spanned by *knotted spin networks*.

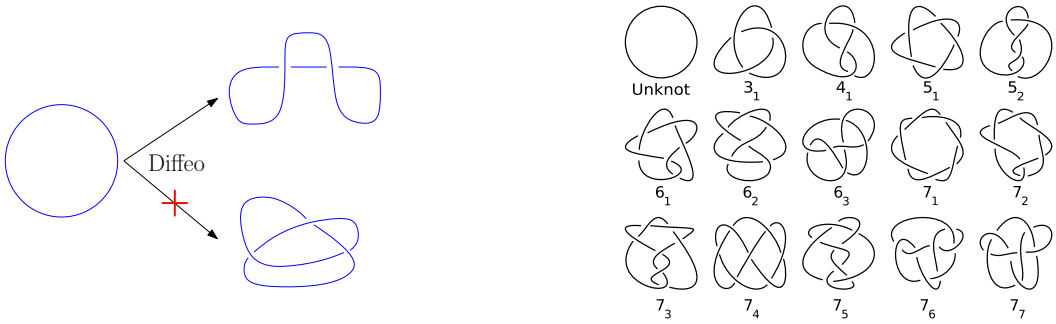


Figure 1: *Left panel.* A diffeomorphism can change the way a graph is embedded in Σ, but not the presence of knots within the graph. *Right panel.* The first few knots (without nodes), taken from Wikipedia.

5.2 The Hamiltonian constraint

Finally, we approach the last step of Dirac's program. On the space of knotted spin networks $\mathcal{H}_{Diff}$, we want to define the Hamiltonian constraint, and study its solutions. The classical scalar constraint is

$$\begin{aligned} H(N) &= \int \mathrm{d}^3x N\epsilon^{ij}_k \frac{E^a_i E^b_j}{\sqrt{\det(E)}} \left(F^k_{ab} - 2\left(1+\gamma^2\right) K^i_{[a} K^j_{b]}\right) \\ &= H^E(N) - 2\left(1+\gamma^2\right) T(N), \end{aligned} \tag{199}$$

where we introduced the shorthand notation $H^E(N)$ and $T(N)$. As with the ADM Hamiltonian constraint (19), this expression is non-linear, which anticipates difficulties to turn it into an operator. However, a trick due to Thiemann [43] allows us to rewrite (199) in a way amenable to quantization. Denoting $V = \int \sqrt{\det(E)}$ the volume of Σ, and

$$\bar{K} = \int K^i_a E^a_i, \tag{200}$$

we can use the classical brackets (106) to establish the following identities,

$$K^i_a = \frac{1}{\gamma}\left(A^i_a - \Gamma^i_a(E)\right) = \frac{1}{\gamma}\left\{A^i_a, \bar{K}\right\}, \tag{201}$$

$$\bar{K} = \frac{1}{\gamma^{3/2}}\left\{\mathcal{H}^E(N\equiv 1), V\right\}, \tag{202}$$

$$\frac{E^a_i E^b_j}{\sqrt{\det(E)}}\epsilon^{ijk}\epsilon_{abc} = \frac{4}{\gamma}\left\{A^k_a, V\right\}. \tag{203}$$

Using these relations,

$$H^E(N) = \int \mathrm{d}^3x\ N\epsilon^{abc}\delta_{ij}F^i_{ab}\left\{A^j_c, V\right\}, \tag{204}$$

$$T(N) = \int \mathrm{d}^3x\ \frac{N}{\gamma^3}\epsilon^{abc}\epsilon_{ijk}\left\{A^i_a, \left\{H^E(1), V\right\}\right\}\left\{A^j_b, \left\{H^E(1), V\right\}\right\}\left\{A^k_c, V\right\}.$$

The advantage of this reformulation is that the non-linearity is mapped into Poisson brackets. The next step is to rewrite these expressions in terms of holonomies and fluxes, so that we can turn them into operators. Notice that we already know the volume operator, and its spectrum can be explicitly computed. This is very promising towards the prospect of knowing the action of the Hamiltonian

constraint. Next, the connection and curvature have to be written in terms of holonomies. This requires a regularization procedure. We describe it only for the first term, H^E, and refer the reader to the literature [24] for $T(N)$.

The connection can be easily expressed in terms of holonomies. Writing explicitly (110), we have that for a path e_a of length ϵ along the x^a coordinate, $h_{e_a}[A] \simeq 1 + \epsilon A_a^i \tau_i + O(\epsilon^2)$, therefore we also have

$$h_{e_a}^{-1} \{h_{e_a}, V\} = \epsilon \left\{A_a^i, V\right\} + O(\epsilon^2). \tag{205}$$

For the curvature, consider an infinitesimal triangular loop α_{ab}, lying on the plane ab, and with coordinate area ϵ^2. At lowest order in ϵ,

$$h_{\alpha_{ab}} = 1 + \frac{1}{2}\epsilon^2 F_{ab}^i \tau^i + O(\epsilon^4), \tag{206}$$

thus

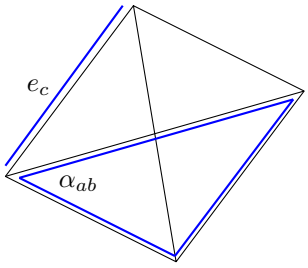
$$h_{\alpha_{ab}} - h_{\alpha_{ab}}^{-1} = \epsilon^2 F_{ab}^i \tau^i + O(\epsilon^4). \tag{207}$$

At this point, we proceed as we did for the geometric operators in the previous section. We introduce a cellular decomposition of Σ, and regularize the integral as a Riemann sum over the cells C_I,

$$H^E = \lim_{\epsilon \to 0} \sum_I \epsilon_I^N 3\epsilon^{abc} \mathrm{Tr}\left(F_{ab} \{A_c, V\}\right) = \tag{208}$$

$$= \lim_{\epsilon \to 0} \sum_I N_I \epsilon^{abc} \mathrm{Tr}\left(\left(h_{\alpha_{ab}} - h_{\alpha_{ab}}^{-1}\right) h_{e_c}^{-1} \{h_{e_c}, V\}\right). \tag{209}$$

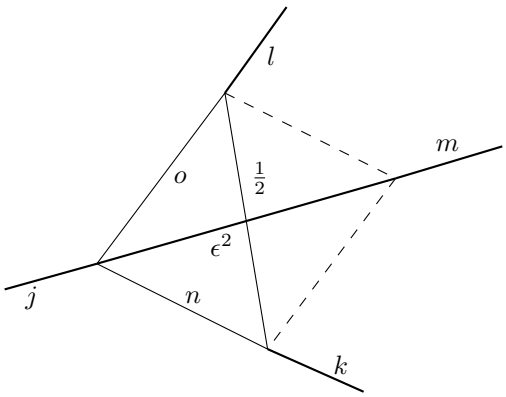
This time is more convenient to specify the cellular decomposition in terms of a *triangulation*, namely a collection of tetrahedral cells. The loop α_{ab} can then be adapted to the triangular faces of this decomposition, as in the following figure.



This expression can now be promoted to an operator in the quantum theory,

$$\widehat{H}^E = \lim_{\epsilon \to 0} \sum_I N_I \epsilon^{abc} \mathrm{Tr}\left(\left(\hat{h}_{\alpha_{ab}} - \hat{h}_{\alpha_{ab}}^{-1}\right) \hat{h}_{e_c}^{-1} \left[\hat{h}_{e_c}, \hat{V}\right]\right). \tag{210}$$

This is a well-defined operator, whose action is explicitly known. It inherits the property of the volume operator of acting only on the nodes of the spin network. From the holonomies, it modifies the spin network by creating new links carrying spin 1/2 around the node, see the following figure.



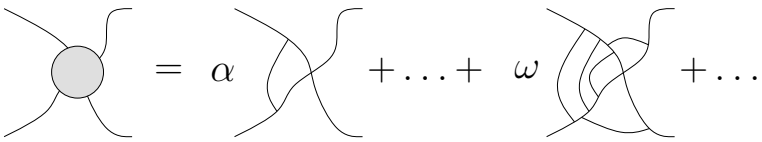
Finally, its amplitude depends on the details of the action of the volume operator. See the literature [22, 24, 40] for details.

Since the only dependence on ϵ is on the position of the new link, in the Hilbert space $\mathcal{H}_{Diff}$ where the position of the link is irrelevant, the limit can be safely taken without affecting the result. Not only this operator is well-defined, also the Dirac algebra (23) can be realized at the quantum level. To see this, we need first to note that the new "exceptional" links added by (210) carry zero volume and are thus invisible to a further action by (210). This crucial property requires the Ashtekar-Lewandowski version of the volume quantum operator, because the new nodes at the junction of the new link with the old ones are planar, and the AL volume operator vanishes on planar nodes. Thanks to this, $\widehat{\mathcal{H}}(N_1)$ and $\widehat{\mathcal{H}}(N_2)$ commute on the space $\mathcal{H}_{Diff}$ of diffeomorphic-invariant spin networks. Hence for diffeomorphism invariant states $\langle\phi| \in \mathcal{H}_{Diff}$,

$$\langle\phi| \left[\widehat{\mathcal{H}}(N_1), \widehat{\mathcal{H}}(N_2)\right] |\psi\rangle = 0 = \langle\phi| \mathcal{H}(N_{[1}\vec{\nabla}N_{2]}) |\psi\rangle\,, \qquad \forall\psi \in \mathcal{H}_{kin}. \tag{211}$$

This is the correct commutator algebra on-shell. Thanks to this non-trivial property, the quantization is said to be anomaly-free.

Another key result is that an infinite number of states solutions of $\widehat{H}$ are known: any graph without nodes is in the kernel of $\widehat{H}^E$ and $\widehat{T}$. This result [12], together with the discreteness of the spectra [11], started the whole interest in the loop quantum gravity approach. More in general, a formal solution of the Hamiltonian constraint will be a particular linear combination (generally infinite) of spin networks with an arbitrary number of exceptional links, whose coefficients depend on the details of the quantum operator and on the spins carried by the spin network,



This construction defines a new spin network with a "dressed node" as the solution. However, this procedure is only formal, and no explicit solutions are known in general.

In conclusion, we have a perfectly well-defined Hamiltonian constraint, whose action is explicitly known and *finite*. An infinite number of states solutions of $\widehat{H}$ are known, and the Dirac algebra is anomaly-free on physical states. Compare this with the old-fashioned Wheeler-De Witt equation (37), which was badly ill-defined, and we see the full force of the use of the Ashtekar variables to quantize general relativity.

Nevertheless, the program of quantization is far from being complete: the complete characterization of $\mathcal{H}_{phys}$ nor the full spectrum of $\widehat{H}$ are known. In spite of the successes of this approach, some limitations

need to be stressed. In particular, although well-defined, the Hamiltonian constraint is plagued by a number of ambiguities:

- We can change the spin of the exceptional edges to an arbitrary value j. This change have significant physical effects in loop quantum cosmology for example.
- We can regularize the connection and the curvature with different paths and still make a consistent theory [40]. In particular, one can envisage a different construction in which the constraint acts on more than one node simultaneously, as advocated in [44].
- Alternative orderings can be explored.

5.3 Current approaches

The effort to gain control over these ambiguities has flourished into two main lines of research. The first one is the idea of the Master Constraint. There one defines a unique constraint implementing simultaneously the diffeomorphisms and scalar constraints. The second one is the spin foam formalism. There one abandons the canonical approach, and seeks a functional integral description of transition amplitudes between spin network states. These two approaches have seen recent important developments, and are the current state of the art of the field.

In the next chapter we will review the spin foam formalism but not the Master Constraint approach, for a review on this one see references [24, 45].

To conclude, although the kinematics of loop quantum gravity is beautifully under control, the dynamics is still work in progress.

We remark nonetheless that the key result on space discreteness, derived using kinematical states, is expected to hold also for physical states [22].

6 Spin Foam

The spinfoam formalism [56, 57, 59, 60, 61, 62] can ideally be viewed as the covariant version of canonical loop quantum gravity (LQG) analyzed in previous chapters. This scenario is nicely realized in three dimensions [63], and there are recent attempts to implement it in quantum cosmology [64, 65]. An important step ahead towards the realization of this scenario in the complete four dimensional theory has been taken with the recent introduction of two strictly related spin-foam models whose kinematics appears to match the one of LQG rather well, which we refer to as the new model [66, 67, 68, 69] and the Freidel-Krasnov-Livine-Speziale (FKLS) model [70, 71].closure

6.1 Preliminary operations

The starting point is the Holst action for general relativity

$$S_{\rm H}(e,A) \;=\; \frac{1}{2}\int_{\Sigma\times\mathbb{R}} \frac{1}{2}\epsilon_{IJKL}e^I\wedge e^J\wedge F^{KL} \;+\; \frac{1}{\gamma}e^I\wedge e^J\wedge F_{IJ}, \tag{212}$$

The theory can be interpreted as given by the action

$$S(E,A) \;=\; \int_{\Sigma\times\mathbb{R}} J_{IJ}\wedge F^{IJ}, \tag{213}$$

with additional constraints

$$J_{IJ} \;=\; \frac{1}{2}\epsilon_{IJKL}e^K\wedge e^L \;+\; \frac{1}{2\gamma}e_I\wedge e_J, \tag{214}$$

In order to obtain the physical Hilbert space we quantize 213 and than we impose the constraints 214.

6.2 The physical Hilbert space

Consider a fixed 4-dimensional triangulation Δ, which is formed by oriented 4-simplices, tetrahedra, triangles, segments and points[6]. The cellular complex Δ^* dual to this triangulation Δ, is made by faces f, edges e and vertices v, dual respectively to triangles f, tetrahedra t and 4-simplices v of Δ.

Given a 3-surface Σ intersecting no vertices of Δ^*, let $\Gamma\equiv\Delta^*\cap\Sigma$. We start from the Hilbert space associated with Σ [68, 69]:

$$H_\Sigma = L^2\left(SL(2,\mathbb{C})^L, \mathrm{d}\mu\right), \tag{215}$$

where μ is the Haar measure on the group $SL(2,\mathbb{C})$, L denotes the number of links in Γ. We fix the orientation such that the node $n=e\cap\Sigma$ is the source of the link $l=f\cap\Sigma$.

By Peter-Weyl theorem, H_Σ can be decomposed as follows

$$H_\Sigma = \bigoplus_{\chi_l}\bigotimes_{l}\left(H^*_{\chi_l}\otimes H_{\chi_l}\right), \tag{216}$$

where χ_l is an assignment of an $SL(2,\mathbb{C})$ representation to each link l and H_χ is the carrier space of the representation χ. The two Hilbert spaces associated to the link l are naturally associated to the two nodes that bound the link l, because they transform under the action of a gauge transformation at one end of the link. Regrouping the four Hilbert spaces associated to each node n, the last equation can be

[6] As we will see the boundary states of this spin foam model is not an arbitrary state of LQG, but only a state with 4-valent intertwiner. This restriction is due to the difficulties in dealing with a non simplicial cellation: a priori the cells are not rigid, is not clear how to impose gluing constraints with cells of different shape. Our contribution that is discussed in the second part of this thesis will help in clearing these difficulties.

rewritten in the form

$$H_\Sigma = \bigoplus_{\chi_l} \bigotimes_n H_n, \qquad H_n = \bigotimes_{a=1}^{4} H_{\chi_a}, \tag{217}$$

where H_n is the Hilbert space associated to a node and a runs here over the four links that join at the node n (that is, the four faces of the boundary tetrahedron t), and we have identified the Hilbert space carrying a representation and its dual.

Here the nodes n label the tetrahedra t in the boundary. We restrict our attention to a single boundary tetrahedron, and its associated Hilbert space H_n, which we call simply H in the following.

Consider the irreducible unitary representations of the principal series of $SL(2,\mathbb{C})$ (for details see [74, 75]), H_n has the structure

$$H = \bigotimes_{a=1}^{4} H_{(k_a,p_a)}, \tag{218}$$

with k a nonnegative integer and p real. The physical intertwiner state space $\mathcal{K}_{\rm ph}$ is a subspace of this space, where the constraints hold in a suitable sense.

As a first step to give the physical boundary space, let us restrict the representations to the ones that satisfy [72]

$$p = \gamma(k+1) \tag{219}$$

We call γ-simple the $SL(2,\mathbb{C})$ representations that satisfy this relation. With this relation, the continuous label p becomes quantized, because k is discrete. It is because of this fact that any continuous spectrum depending on p comes out effectively discrete on the subspace satisfying the relation (219).

Next, fix an $SU(2)$ subgroup of $SL(2,\mathbb{C})$, then the (k,p) representation for the single component of H associated with a single boundary face f splits into the irreducible representations H_j of the $SU(2)$ subgroup as

$$\mathcal{H}_{(k,p)} = \bigoplus_{j=k}^{\infty} \mathcal{H}_j, \tag{220}$$

with j increasing in steps of 1. Consider the lowest spin term in each factor, where j in the decomposition (220) is reduced to

$$j = k; \tag{221}$$

this selects the "minimal" subspace

$$\mathcal{H}^{\rm min} = \bigotimes_{a=1}^{4} \mathcal{H}_{k_a}. \tag{222}$$

The final physical intertwiner space $K_{\rm ph}$ is given by the $SU(2)$-invariant subspace of $\mathcal{H}^{\rm min}$:

$$\mathcal{K}_{\rm ph} = {\rm Inv}_{\rm SU(2)}[\mathcal{H}^{\rm min}]. \tag{223}$$

The total physical boundary space $\mathcal{H}_{\rm ph}$ of the theory is then obtained as the span of spin-networks in $L^2[SL(2,\mathbb{C})^L/SL(2,\mathbb{C})^N]$ with γ-simple representations on edges and with intertwiners in the spaces $\mathcal{K}_{\rm ph}$ at each node. In the next subsection, we will show this physical boundary space $\mathcal{H}_{\rm ph}$ solves all the kinematic constraints in a suitable sense.

6.3 Kinematic constraints

Now let us come to introduce the kinematic constraints, including the simplicity constraints and the closure constraint, and show all of them are satisfied on the physical boundary space $\mathcal{H}_{\rm ph}$. On the fixed oriented triangulation Δ, we restrict the metric to be a Regge one [76]: flat within each 4-simplex, with

curvature on the triangles. We choose as the boundary variables $\mathfrak{sl}(2,\mathbb{C})$-valued variables J_l^{IJ} associated with the links l of the graph formed by the one-skeleton of the cellular complex dual to the boundary triangulation. We also need constraints to restrict these variables to the gravity fields, it can be shown [66, 67, 68] that (214) are equivalent to:

$$\textit{Simplicity constraints}: \qquad C_l^J = n_I \left((*J_l)^{IJ} + \frac{1}{\gamma} J_l^{IJ} \right) = 0, \tag{224}$$

$$\textit{Closure constraints}: \qquad G^{IJ} = \sum_{a=1}^{4} J_{l_a}^{IJ} = 0, \tag{225}$$

where n_I denotes the normal to the tetrahedron t, and $*$ stands for the Hodge dual in the internal indices; the sum is over the four links that join at the node dual to the tetrahedron (or over the four triangles bound the tetrahedron). These constraints will give the solution $B_f = \int_f *\big(e(t) \wedge e(t)\big)$, where $e(t)$ is a tetrad one-form covering the tetrahedron t, and $J_f = B_f + \frac{1}{\gamma} * B_f$, with triangle f dual to the link l. In particular, if we choose a "time" gauge where $n_I = (0,0,0,1)$, the simplicity constraint (224) turns out to be

$$C_l^i = J_l^{0i} + \gamma\, {}^*J_l^{0i} = 0, \tag{226}$$

which leads to the key constraint of the new model

$$C_l^i = K_l^i + \gamma\ L_l^i = 0, \tag{227}$$

where $L_l^i \equiv \frac{1}{2}\epsilon^i{}_{jk} J_l^{jk}$ and $K_l^i \equiv J_l^{0i}$ are respectively the generators of the $SU(2)$ subgroup that leaves n_I invariant, and the generators of the corresponding boosts. In terms of these generators, the closure constraint (225) becomes

$$G_L^i = \sum_{a=1}^{4} L_l^i = 0 \qquad\qquad G_K^i = \sum_{a=1}^{4} K_l^i = 0. \tag{228}$$

To quantize the constraints (227) (228), one just need to replace the generators with the associated operators.

Generators on the canonical bases. Given a carrier space $H_{(k,p)}$, the canonical basis is given by the basis diagonalizing simultaneously the Casimir operators $J\cdot J$, $*J\cdot J$, $L\cdot L$ and L^3 , which is noted as $|(k,p);j,m\rangle$ or simply as $|j,m\rangle$. On this canonical basis, the generators act in the following way:

$$\begin{aligned}
L^3|j,m\rangle =& m|j,m\rangle, \\
L^+|j,m\rangle =& \sqrt{(j+m+1)(j-m)}|j,m+1\rangle, \\
L^-|j,m\rangle =& \sqrt{(j+m)(j-m+1)}|j,m-1\rangle, \\
K^3|j,m\rangle =& -\alpha_{(j)}\sqrt{j^2-m^2}|j-1,m\rangle - \beta_{(j)} m|j,m\rangle + \alpha_{(j+1)}\sqrt{(j+1)^2-m^2}|j+1,m\rangle, \\
K^+|j,m\rangle =& -\alpha_{(j)}\sqrt{(j-m)(j-m-1)}|j-1,m+1\rangle - \beta_{(j)}\sqrt{(j-m)(j+m+1)}|j,m+1\rangle \\
& -\alpha_{(j+1)}\sqrt{(j+m+1)(j+m+2)}|j+1,m+1\rangle, \\
K^-|j,m\rangle =& \alpha_{(j)}\sqrt{(j+m)(j+m-1)}|j-1,m-1\rangle - \beta_{(j)}\sqrt{(j+m)(j-m+1)}|j,m-1\rangle \\
& +\alpha_{(j+1)}\sqrt{(j-m+1)(j-m+2)}|j+1,m-1\rangle,
\end{aligned} \tag{229}$$

where

$$L^{\pm} = L^1 \pm iL^2, \qquad K^{\pm} = K^1 \pm iK^2,$$
$$\alpha_{(j)} = \frac{i}{j}\sqrt{\frac{(j^2-k^2)(j^2+p^2)}{4j^2-1}}, \qquad \beta_{(j)} = \frac{kp}{j(j+1)}.$$

Now let us go to show the physical Hilbert space $\mathcal{H}_{\rm ph}$ derived last subsection solves indeed the constraint operators associated to the simplicity constraints (227) and the closure constraints (228). Namely, we will show

(i) the simplicity constraints (227) are satisfied in the "minimal" γ-simple representation $\mathcal{H}^{\rm min}$,

(ii) the closure constraints (228) are satisfied in the intertwiner space $\mathcal{K}_{\rm ph}$.

To show (i), let us consider the states in the "minimal" space $\mathcal{H}^{\rm min}$ in equation (222). For these lowest spin states, equation (221) implies that the states are of the form $|(k,p);k,m\rangle$, or simply as $|k,m\rangle$. The action (229) of the generators on these states reads:

$$\begin{aligned}
\big(K^3+\beta_{(k)}L^3\big)|k,m\rangle &= \alpha_{(k+1)}\sqrt{(k+1)^2-m^2}|(k+1,m)\rangle,\\
\big(K^+ +\beta_{(k)}L^+\big)|k,m\rangle &= -\alpha_{(k+1)}\sqrt{(k+m+1)(k+m+2)}|(k+1,m+1)\rangle,\\
\big(K^- +\beta_{(k)}L^-\big)|k,m\rangle &= \alpha_{(k+1)}\sqrt{(k-m+1)(k-m+2)}|(k+1,m-1)\rangle.
\end{aligned}$$

It is straightforward to obtain

$$\langle k,m'|\big(K^i+\beta_{(k)}L^i\big)|k,m\rangle = 0. \tag{230}$$

Using the relation (219), $\beta_{(k)}$ turns out to be the Barbero-Immirzi parameter γ and the matrix elements of the l.h.s of (227) hence vanish on the "minimal" γ-simple space:

$$\langle k,m'|C^i|k,m\rangle = \langle k,m'|\big(K^i+\gamma L^i\big)|k,m\rangle = 0. \tag{231}$$

Notice that the slight difference of our relation (227) from the old one plays a key role here. Notice also that what we obtain is that the matrix elements vanish *exactly*, and not just in the large spin limit.

To show (ii), observe that the l.h.s. of the right equation of (228) is the generator of $SU(2)$ transformations at the node and vanishes strongly on (223) by definition; the l.h.s. of the left equation of (228) is proportional to the one the right equation of (228) by (231) and therefore vanishes weakly. Thus $\mathcal{K}_{\rm ph}$ is the intertwiner space as a solution of *all* the constraints: all the constraints hold weakly.

Summarizing, we have introduced the kinematic constraints and shown that all of them are satisfied on the physical boundary space $\mathcal{H}_{\rm ph}$ derived in the last subsection. Since we have not proven that the physical Hilbert space considered is the *maximal* space where the constraints hold weakly, one might worry that the physically correct quantization of the degrees of freedom of general relativity could need a larger space. Also, it has been pointed out that imposing second class constraints weakly might lead to inconsistencies in some cases [72]. In the present case, however, these worries are not relevant, since the space obtained is directly related to the one of the canonical theory, which we can trust to capture the degrees of freedom of gravity correctly.

6.4 Definition of the amplitude

In this subsection we want associate an amplitude W to each transition from a spin-network state to another on the boundary of a spin foam model. This amplitude is largely determined by general principles:

superposition, locality, diffeomorphism invariance, crossing symmetries and Lorentz invariance. The *superposition* principle require that the amplitude W can be expanded in a sum over "histories of states"

$$W = \sum_{\sigma} W(\sigma) \tag{232}$$

where $W(\sigma)$ is an amplitude associated to an appropriated sequence of states σ bounded by the considered spin-network.
Because of the *locality* principle we expect that $W(\sigma)$ decompose in terms of product of elementary amplitudes W_v associated to local elementary processes

$$W(\sigma) \sim \prod_{v} W_v \tag{233}$$

the amplitude W_v will be interpreted as an elementary vertex in the same sense as two electrons one photon vertex is the elementary amplitude of QED.
The *diffeomorphism invariance* requirement is intended in a very loose sense. Because the in the canonical approach th Hamiltonian constraint is a density under diffeomorphism we expect that also in the covariant formulation the evolution operator acts only on nodes of spin networks, and in particular that W_v is associated to processes that transforms the nodes of the boundary states.
We also want that *crossing symmetry* holds. That is the well known property of standard QFT that the vertex amplitude is independent on which states are interpreted as "in" or "out".
The last but more important ingredient is the *Lorentz invariance*. Since classical general relativity has a local Lorentz invariance we expect the individual Spin Foam vertex to be Lorentz invariant in some sense. This enters deeply in the definition of W_v

$$W_v = \int_{SL(2,\mathbb{C})^N} \mathrm{d}\tilde{g}_n \prod_{\ell} \sum_{j} (2j+1) \mathrm{Tr}[D^{(j)}(U_\ell) Y_\gamma^\dagger D^{(\gamma j, j)}(g_{s(\ell)} g_{t(\ell)}^{-1}) Y_\gamma] \tag{234}$$

where $\mathrm{d}\tilde{g}_n = \delta(g_1)\mathrm{d}g_n$ is a regularized $SL(2,\mathbb{C})^N$ measure, the product over ℓ is the product over all the faces that converge in the vertex, $D^{(p,k)}$ is the ρ, k unitary irreducible representation of $SL(2,\mathbb{C})$ and the Y_γ is the injection map from the $SU(2)$ spin j representation to the "minimal" subspace in γ-simple representation (222)

$$Y_\gamma : H_j \to H_j \subset H_{(p=\gamma(j+1), k=j)} \tag{235}$$

By bringing all together we have the full amplitude associated to a boundary spin network state ψ. This can be expressed as the "spinfoam sum"

$$\langle W|\psi\rangle = \sum_{\sigma} \prod_{f} d(j_f) \prod_{v} W_v(\sigma). \tag{236}$$

The sum is over spinfoams σ bounded by the spin network ψ. If we "cut a spinfoam with a 2d-surface", we obtain a spin network: the intersection of the edges e with the surface gives the nodes n of the spin network and the intersection of the faces f with the surface gives the links l of the spin network, with their respective colorings. The face amplitude is the dimension of the $SU(2)$ representation coloring the face $d(j_f) = (2j_f + 1)$[73].

In particular, an S_3 surface surrounding a vertex v of σ defines a spin network ψ_v. The vertex amplitude of the vertex v of σ is defined to be

$$W_v(\sigma) \equiv \langle W_v|\psi_v\rangle. \tag{237}$$

This is "local", in the sense that it depends only on the spins and intertwiners surrounding the vertex.

7 Approaching quantum geometry

Let us review what we have learned in Section 4. The golden tools of loop quantum gravity are the spin network states. These form a basis in the kinematical Hilbert space, and diagonalize geometric operators. In particular, the quantum numbers carried by a spin network, (Γ, j_e, i_n), define a notion of *quantum geometry*,[7] in the same way as the quantum number of an harmonic oscillator defines its quantum state. These quantum numbers can be compared with the kinematical metric g_{ab}, defining the classical geometry of space. We use the word kinematical to mean that g_{ab} is an arbitrary metric, not necessarily a solution of Einstein's equations, just like an arbitrary spin network spans the kinematical Hilbert space, not necessarily solving the diffeomorphisms and Hamiltonian constraints. The quantum geometry described by (Γ, j_e, i_n) is very different than a classical geometry g_{ab}. Specifically, it is largely insufficient to reconstruct a metric g_{ab}. We can highlight three key differences:

(i) Quantized spectra: the spectra of geometric operators are discrete, as opposed to the continuum values of their classical counterparts. This is a standard situation in quantum mechanics, think for instance of the discretization of the energy levels of an harmonic oscillator.

(ii) Non-commutativity: not all geometric operators commute among themselves. This is a consequence of the non-commutativity of the fluxes (148). This is also standard, like the incompatibility of position and momentum observables.

(iii) Distributional nature: the states capture only a finite number of components of the original fields, that is their values along paths (for the connection) and surfaces (for the triad). This is reminiscent of what happens in lattice theories, where the continuum field theory is discretized on a fixed lattice and only a finite number of degrees of freedom are captured.

In spite of these differences, if the theory is correct it must admit a *semiclassical* regime where a smooth geometry emerges. And furthermore, the dynamics of this smooth geometry should be given by general relativity, at least in some approximation. This is what we expect from any quantum theory: that the $\hbar \mapsto 0$ limit is well-defined, and reproduces classical physics. The points (i-iii) show that this limit might be non-trivial to obtain in loop quantum gravity. Although we started from genuine general relativity and quantum mechanics, without any exotic ingredient, we ended up with a description in terms of quantities, (Γ, j_e, i_n), which are very far from the original ones, g_{ab}. This is the problem of the semiclassical limit.

In this Section, we would like to comment on the first part of this problem, namely the emergence of a smooth geometry. This requires understanding the points (i-iii). The first point is easier to deal with: also the orbitals of the hydrogen atoms are quantized, placed at distances labeled by an integer n. The classical Keplerian behavior is recovered if we look at the large n limit. Similarly, continuum spectra are recovered in the large spin limit $j_l \mapsto \infty$. Points (ii) and (iii) are more subtle. The key to deal with them is the use of *coherent states*, namely linear superpositions of spin network states peaked on a smooth geometry.

Recall that coherent states are peaked on a point in the classical phase space. For instance, the coherent state $|z\rangle$ for an harmonic oscillator is peaked on the initial position $q = \mathrm{Re}(z)$ and momentum $p = \mathrm{Im}(z)$. The phase space is the familiar cotangent bundle to the real line, $(q, p) \in T^*R$. Analogously,

[7] Each face dual to a link has an area proportional to the spin j_e, and each region around a node has a volume determined by the intertwiner i_n as well as the spins of the link sharing the node.

coherent states for loop quantum gravity are peaked on a point $(A_a^i(x), E_i^a(x))$ in the classical phase space, which defines an intrinsic (through the triad) and extrinsic (through the connection) 3-geometry. Such coherent states for loop quantum gravity were introduced in [47]. These states minimize the uncertainty of the flux operators, thus addressing (ii). In order to address (iii) and recover a smooth geometry *everywhere* on Σ, the coherent states have support over an infinite number of graphs.

On the other hand, dealing with an infinite number of graphs is a formidable task, and for practical purposes, one often needs to rely on approximations. A convenient one is to allow only states living on a fixed finite graph Γ. The Hilbert space $\mathcal{H}_\Gamma$ provides a *truncation* of the theory, which may be sufficient to capture the physics of appropriate regimes. Within this truncation, one considers coherent states in $\mathcal{H}_\Gamma$. However, in which sense one can assign a classical geometrical interpretation to these states? They will be peaked on points in the phase space corresponding to $\mathcal{H}^0_\Gamma$, which consists of classical holonomies and fluxes on Γ. These quantities capture only a *finite* number of components of a continuum geometry. To be able to interpret these truncated coherent states in a physical sense, we need to use these data to approximate a continuum geometry. The problem is similar to approximating a continuous function $f(x)$ if we are given a finite number of its values, $f_n = f(x_n)$. Various interpolating procedures are common. We now describe a notion of interpolating geometry which emerges naturally. What follows is in part an anticipation of the next sections where the full details are given. Now we do a summary that we think is useful in order not to lose what's our goal in the following sections.

The associated gauge-invariant Hilbert space on a fixed graph is a sum over intertwiner spaces $\mathcal{H}_n \equiv \mathrm{Inv}[\bigotimes_{e\in n} V^{(j_e)}]$ associated to the nodes (see (169)),

$$\mathcal{H}^0_\Gamma = \bigoplus_{j_l} \bigotimes_n \mathcal{H}_n. \tag{238}$$

Just like $L_2(\mathbb{R}, \mathrm{d}x)$ is the quantization of the classical phase space T^*R, this Hilbert space is the quantization of a classical phase space[8], denote it S_Γ. The important result which is relevant for us is that this phase space can be parametrized as follows [48],

$$S_\Gamma = \bigtimes_l T^*S^1 \times \bigtimes_n \mathcal{S}_{F(n)}. \tag{239}$$

Here T^*S^1 is the cotangent bundle to a circle, F the valency of the node n, and $\mathcal{S}_F$ is a $2(F-3)$-dimensional phase space, introduced by Kapovich and Millson [49]. This parametrization mimics the decomposition (238). In particular, the quantization of T^*S^1 gives the quantum number j_l associated to each link, and the quantization of $\mathcal{S}_{F(n)}$ gives the $F-3$ intertwiners associated to an F-valent node.

The parametrization defines a notion of interpolating geometry associated to a cellular decomposition of Σ dual to Γ, called *twisted geometry* [48]. The key to the geometric interpretation of (239) is the node space $\mathcal{S}_{F(n)}$. Fix the spins j_l on the links connected to the node, and consider F unit vectors n_l in $\mathbb{R}^3$ constrained to satisfy the following *closure condition*,

$$C = \sum_{l=1}^{F} j_l n_l = 0. \tag{240}$$

Because of the closure, the vectors $j_l n_l$ span a polygon in $\mathbb{R}^3$, see Figure 2. The space of all polygons with fixed j_l up to SO(3) rotations forms the phase space $\mathcal{S}_F$, called by Kapovich and Millson the space of shapes of the polygon.

[8] Up to singular points, but this is not important for the following.

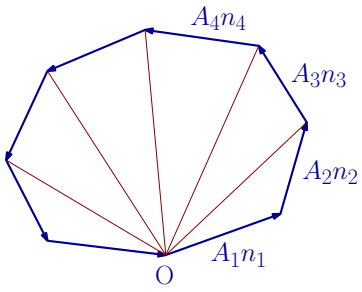


Figure 2: A polygon in $\mathbb{R}^3$, with F sides of fixed lengths. Because of the closure condition, not all the components of the normals are independent, but only $2F-3$. If we further consider the polygon up to rotations, that is the space $\mathcal{S}_F$, this is only specified by $2F-3-3=2(F-3)$ numbers. These can be taken to be the lengths of $F-3$ diagonals and their conjugated dihedral angles (which measure the way the polygon hinges on each diagonal).

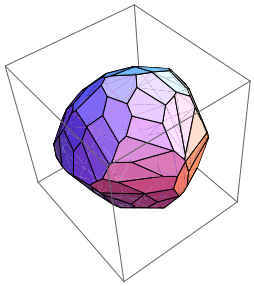


Figure 3: A generic polytope with 100 faces. Notice that most faces are hexagons.

This means that we can interpret the intertwiners as "quantized polygons". Such interpretation is appealing, but not particularly useful for our purposes. More promising would be a description in terms of *polyhedra*, rather than polygons: polyhedra can be glued together to approximate a smooth manifold, and their geometry will induce a discrete metric of some sort.

The good news [51] is that this is precisely what happens! In fact, each (non-coplanar) configuration of vectors n_l describes also a *unique polyhedron* in $\mathbb{R}^3$, with j_l as the areas of the faces, and n_l as their unit normals. More precisely, a convex bounded polyhedron, which is usually referred to as a "polytope". The explicit reconstruction of the polytope geometry from holonomies and fluxes is studied in [51]. Therefore, up to the degenerate configurations corresponding to coplanar normals, we can visualize $\mathcal{S}_F$ as a polytope with F faces.

Going back to (239), the T^*S^1 space on the links carries information needed to define the parallel transport between adjacent polytopes. This unique correspondence between $\mathcal{S}_F$ in (239) and polytopes means that *each (non-degenerate) classical holonomy-flux configuration on a fixed graph can be visualized as a collection of adjacent polytopes, each one with its own frame, and with a notion of parallel transport between any two polytopes.* Since twisted geometries parametrize the classical phase space on a fixed graph, coherent states on a fixed graph are peaked on the approximated smooth geometry described by a twisted geometry, i.e. a collection of polytopes.

Twisted geometries have a peculiar characteristic which justifies their name: they define a metric which is locally flat, but *discontinuous*. To understand this point, consider the link shared by two nodes. Its dual face has area proportional to j_l. However, its shape is determined independently by the data around each node (the normals and the other areas), thus generic configurations will give different shapes.

In other words, the reconstruction of two polytopes from holonomies and fluxes does not guarantee that the shapes of shared faces match. Hence, the metric of twisted geometries is discontinuous in the sense that the shape of a face "jumps" when going from one polytope to the next. See left panel of Figure 12.

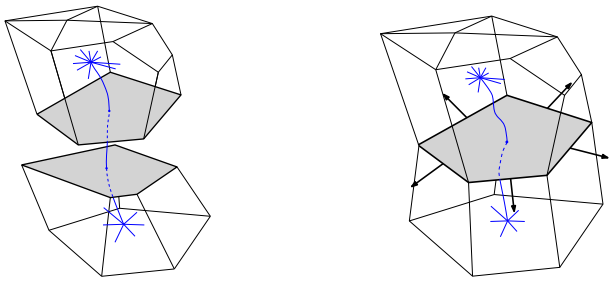


Figure 4: Two high-valency nodes and the link connecting them. Left panel: In general, the two adjacent polytopes, defined by holonomies and fluxes on the nodes, don't glue well. Even if the area of a polygonal face is the same because there is a unique spin associated with the link, the shapes will be different in general. Right panel: in order for the shapes to match, one needs to impose appropriate conditions on the polytopes such as the matching of the normals to the edge in the plane of the face. These conditions, studied in [51], affect the global shape of the polytope.

Notice that one can also consider a special set of configurations for which the shapes match, see right panel of Figure 4. This is a subset of S_Γ which corresponds to piecewise flat and continuous metrics. For the special case in which all the polytopes are tetrahedra, this is the set-up of Regge calculus, and those holonomies and fluxes describe a 3d Regge geometry. The matching conditions in this case were studied in [52]. For an arbitrary graph, the matching conditions are studied in [51], and the result would be Regge calculus on generic cellular decompositions.

The relation between twisted geometry and Regge calculus implies that holonomies and fluxes carry *more* information than the phase space of Regge calculus. We stress that this is not in contradiction with the fact that the Regge variables and the LQG variables on a fixed graph both provide a truncation of general relativity: simply, they define two distinct truncations of the full theory. See [53] for a discussion of these aspects.

Summarizing, when we truncate the theory to a single graph, we capture only a finite number of degrees of freedom of the geometry. This finite amount of information can be suitably interpolated to give an approximate description of a smooth geometry. The reparametrization in terms of twisted geometries achieves a prescription for this interpolation procedure: starting from holonomies and fluxes on a graph, we can assign to them a specific twisted geometry, that is, a bunch of polytopes stuck together.

These results offers on the one hand a new way to visualize the quantum geometry of loop quantum gravity in terms of fuzzy polytopes. On the other hand, they open the door to new beautiful mathematical ingredients, from the geometry of polytopes [50, 51], to twistors [54], to new emerging symmetries [55].

8 Polytopes: definition and combinatorial classes

A polytope is a bounded convex polyhedron, namely the convex hull of a finite set of points. It can be represented as the intersection of finitely many half-spaces as

$$\mathcal{P} = \left\{ x \in \mathbb{R}^3 \,|\, n_i \cdot x \leq h_i, \quad i = 1, \ldots, m \right\}, \tag{241}$$

where n_i are m arbitrary vectors, and $h_i \in \mathbb{R}^m$. This definition is redundant: the minimal set of inequalities corresponds to taking $m = F$ the number of faces of the polytope. It is also abstract: the polytope is characterized by a set of m arbitrary vectors n_i and real numbers h_i. We are interested in a more geometric characterization in terms of a specific set of vectors and numbers: the unit normals and areas of its faces. Consider then a set of unit vectors $n_i \in \mathbb{R}^3$ and a set of positive real numbers A_i such that

$$C \equiv \sum_{i=1}^{F} A_i \, n_i = 0. \tag{242}$$

Each choice of areas and normals defines a polytope (241) as follows. For each vector consider the plane orthogonal to n_i at a distance h_i from the origin of $\mathbb{R}^3$. The intersections of the half-spaces defines the polytope $n_i \cdot x \leq h_i$. We can then adjust the heights to $h_i = h_i(A)$ such that the faces have areas A_i.

Remarkably, to each choice of areas and normals there corresponds a *unique* polytope. This is established thanks to the following theorem (e.g. [50]):

Theorem (Minkowski)

(a) *If $n_1, \ldots, n_F$ are non-coplanar unit vectors and $A_1, \ldots, A_F$ are positive numbers such that* (242) *holds, than there exist a polytope whose faces have outwards normals n_i and areas A_i.*

(b) *If each face of a polytope is equal in area to the corresponding face with parallel external normal of a second polytope and conversely, then the two polytopes are congruent by translation.*

This uniqueness will play an important role in the following. [9]

8.1 Phase space structure

Let us put the polytopes aside for a moment, and consider the following space of closed vectors at fixed norms,

$$\mathcal{P}_F = \left\{ n_i \in (S^2)^F \,|\, C(A_i, n_i) = 0 \right\}. \tag{243}$$

The areas are seen as fixed external parameters. For each configuration of areas, the space has dimensions $2F - 3$. Points in $\mathcal{P}_F$ have a very simple geometric interpretation: the condition (242) makes the vectors close on themselves, thus each configuration defines a polygon in $\mathbb{R}^3$ with side vectors $A_i n_i$ (see Fig. 5). We introduce the *space of shapes*, $\mathcal{S}_F$, of the polygon, as the space of closed unit normals modulo rotations,

$$\mathcal{S}_F \equiv \mathcal{P}_F / \mathrm{SO}(3). \tag{244}$$

Each point in this space represents a polygon in $\mathbb{R}^3$ up to rotations. This $2(F-3)$-dimensional space was shown by Kapovich and Millson [49] to be a symplectic manifold, up to a finite number of points

[9] The key property of this space is that to each (non-degenerate) point there corresponds a *unique* polytope. This is established thanks to the following (e.g. [50]) Thanks to this theorem, a set of (non-coplanar) closed normals and areas defines a unique polytope embedded in $\mathbb{R}^3$. The space $\mathcal{P}_F$ contains also configurations where the vectors are coplanar, which we dub degenerate since they only span a two dimensional space. These degenerate configurations do not define a polytope, but rather a *polygon*. In fact, notice that (242) defines also a unique polygon in $\mathbb{R}^3$, with $A_i n_i$ as the side vectors (see Fig.5). This definition is valid for any point in $\mathcal{P}_F$. For the dense subset of non-coplanar normals, Minkowski's theorem allows one to interpret the configuration also as a specific polyhedron.

corresponding to configurations with one or more flat angles among the sides of the polygon. Action-angle variables for (244) are $(F-3)$ pairs (μ_i, θ_i) with canonical Poisson brackets, $\{\mu_i, \theta_j\} = \delta_{ij}$. Here μ_i is the length of i-th diagonal vector $\vec{\mu}_i = A_1 n_1 + \ldots + A_{i+1} n_{i+1}$ (see Fig.5), and its conjugate variable θ_i is a dihedral angle which measures the way the polygon hinges on the i-th diagonal. Because at fixed areas the range of each μ_i is finite, this phase space has the topology of the Cartesian product of $(F-3)$ cylinders, whose heights depend on A_i.

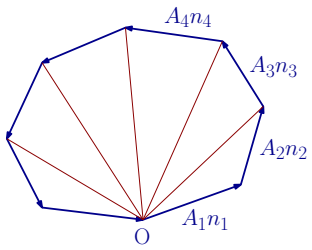


Figure 5: A polygon with side vectors $A_i n_i$ and the $(F-3)$ independent diagonals. The space of possible polygons in $\mathbb{R}^3$ up to rotations is a $2(F-3)$-dimensional phase space, with action-angle variables the pairs (μ_i, θ_i) of the diagonal lengths and dihedral angles. For non-coplanar normals, the same data defines also a unique polytope thanks to Minkowski's theorem.

Thanks to Minkowski's theorem, we know that each point of non-coplanar normals represents also a unique polytope, thus we can consistently identify polygons in $\mathbb{R}^3$ with polytopes. Let us call degenerate the configurations with coplanar normals, since they only span a two dimensional space. Hence, we will refer to (244) as the *space of shapes of polytopes at fixed areas*, with the caveat of the degenerate configurations in mind.[10]

8.2 Classes of polytopes

We are interested in a classification of $\mathcal{S}_F$, thus we study polytopes as we keep the areas fixed, and vary the normals. An important observation is the following: as we vary the normals, not only the intrinsic geometry, but in general also the combinatorial structure of the polytope changes, that is the number of boundary vertices and the valency of the polygonal faces. We refer to the combinatorial structure as the *class* of the polytope. In other words, there are two components to the shape of a polytope: its class, and the intrinsic shape, namely the intrinsic geometry (up to rotations) once the combinatorial structure is fixed. Which classes is realized, depends on the specific values of the normals. This is a point we would like to stress: one is not free to choose a class, and then assign the data. It is on the contrary *the choice of data to select the class*. This is an immediate consequence of Minkowski's theorem.

A useful tool to visualize polytopes and their classes are Schlegel diagrams [85, 50]. The Schlegel diagram of a polytope is a planar graph obtained choosing a face f, and projecting all the other faces on f as viewed from above. See Fig.6 for examples.

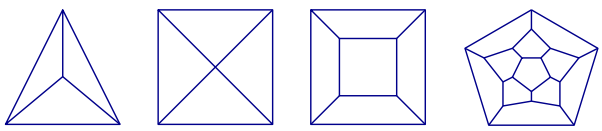


Figure 6: Some examples of Schlegel diagrams. From left to right, a tetrahedron, pyramid, cube and dodecahedron.

[10] Concerning the sick configurations excluded from the genuine phase space, notice that they correspond to redundant description of a polyhedron with lesser F, thus can be safely excluded from the description of the polytope.

So the question is: given a (non-degenerate) point in $\mathcal{S}_F$, what polytope does it identify? In the case $F = 4$, it is well-known that there is always a unique tetrahedron associated with four closed normals. For $F = 5$ there are two possible classes: a triangular prism, and a pyramid, see Fig. 7. The two classes

Figure 7: Polytopes with 5 faces: the two possible classes are the triangular prism (left panel) and the pyramid (right panel).

have different polygonal faces: Two triangles and four romboids for the triangular prism, four triangles and one romboid for the pyramid. The number of vertices changes as well, six for the prism, five for the pyramid.

Minkowski's theorem guarantees us that the same set (A_i, n_i) cannot be associated to both classes: each point in $\mathcal{S}_5$ corresponds to only one of the two. It is then interesting to map the phase space into regions corresponding to different classes. One might at first think that both classes will generically appear, but this is not the case. In fact, notice that the pyramid is just a special case of the prism, obtained collapsing to a point one of the edges connecting two triangular faces. The existence of a pyramid then requires a non-trivial condition, i.e. the presence of a 4-valent vertex. A moment of reflection shows that this condition can be imposed via an algebraic equation on the variables. Hence the shapes corresponding to pyramids span a codimension one surface in $\mathcal{S}_5$. Generic configurations of areas and normals describe triangular prisms, and the pyramids are measure zero special cases. We call *dominant* the class of maximal dimensionality, e.g. the triangular prism here.

Let us move to $F = 6$, a case of particular interest since regular graphs in $\mathbb{R}^3$ are six-valent. There are *seven* different classes of polytopes, see Fig.8. The most familiar one is the cuboid (top left of Fig.8), with its six romboidal faces. It is defined, up to rotations, by 12 numbers (e.g. the edge lengths). Although this is precisely the dimension of $\mathcal{S}_6$, the cuboid class does *not* cover completely, by itself, the space of shapes of six areas and closed normals: there is a second class with 12 edges and 8 three-valent vertices also spanning a 12-dimensional space of intrinsic shapes. It is a "pentagonal wedge" with two triangles,

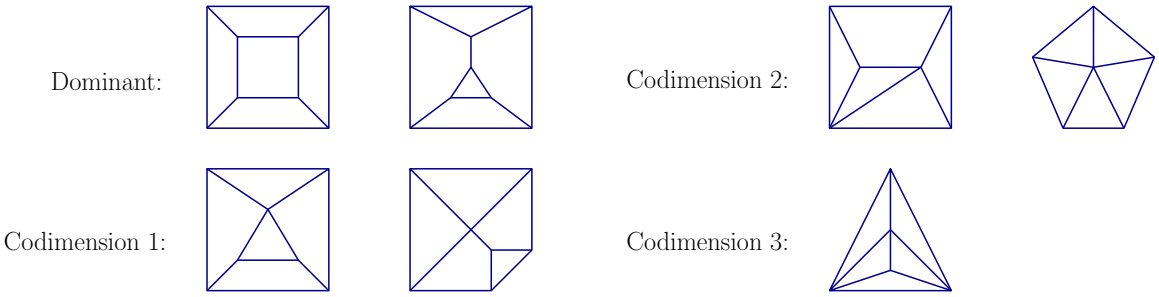


Figure 8: The seven classes of polytopes with 6 faces, grouped according to the dimensionality of their configurations.

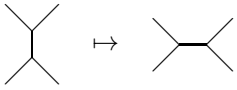
two romboids and two pentagons (to visualize it, imagine a triangular prism planed down on a corner, so that a vertex is replaced by a triangle).

The remaining five classes are subdominant, because non-trivial conditions are required for their existence. The first two are classes with 11 edges and codimension one, which require one 4-valent vertex. These classes can be obtained from the dominant ones for configurations where a specific edge has zero

length, as it can be easily seen from their Schlegel diagrams in Fig.8. From the cuboid, collapsing the upper edge of the inner square gives the codimension-one polytope right below it. From the pentagonal wedge, collapsing one of the two lateral sides gives the polytope right below it. The codimension-two classes have 10 edges, and require either two four-valent vertices (the one on the left), or a single five-valent vertex (the pentagonal pyramid, the one on the right). Again, they can be easily obtained from those of codimension-one setting to zero a specific edge length.

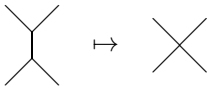
Finally, there is also a class of codimension 3 with 9 edges. It has six triangular faces and three four-valent vertices. This class is interesting in that it can be seen as two tetrahedra glued along a common triangle. The shared triangle is the one with the four-valent vertices. Now, two arbitrary tetrahedra are defined by 12 independent numbers. In order for them to glue consistently and generate this polytope, the shape of the shared triangle has to match. This shape matching requires three conditions (for instance matching of the edge lengths), thus we obtain a 9-dimensional space of shapes. That is, precisely the codimension 3 subspace in $\mathcal{S}_6$ spanned by the class. Hence this class is a special case of two tetrahedra where conditions are imposed for them to glue consistently.

Since all classes correspond to tessellations of the sphere with F faces, they must be connected by Pachner moves [87] not changing the total number of faces. That is, the 2-2 move



(245)

In order to obtain also the subdominant classes, we need to allow edges to collapse. Hence, we consider also the move "frozen" in its intermediate position,



(246)

Using these moves one can connect all classes of polytopes at fixed F. The amused reader can easily find a sequence of (245) and (246) connecting all seven classes of Fig.8. To start, apply the move (245) to the upper edge of the inner square of the cuboid to obtain the pentagonal wedge.

From the above analysis, we expect that the phase space $\mathcal{S}_6$ can then be divided into regions corresponding to the two dominant classes, separated by the subdominant ones. This is qualitatively illustrated in Figure 9. Below in Section 9.4 we will have the right tools to confirm this expectation, and turn the picture into a quantitative one.

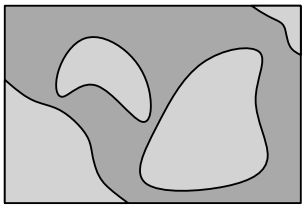


Figure 9: Pictorial representation of the phase space: it can be mapped into regions corresponding to the various dominant classes (two in the example). The subdominant classes separate the dominant ones and span measure-zero subspaces. See Figure 11 below for concrete implementations of this picture.

After this brief survey of some specific examples, let us make some general statements.

- The phase space $\mathcal{S}_F$ can be mapped in regions corresponding to different classes. The dominant classes, generically more than one, cover it densely, whereas the subdominant ones span measure-zero subspaces. Dominant classes in phase space correspond to polytopes with all three-valent

vertices. This condition maximizes both the number of vertices, $E = 2(F-2)$, and edges, $V = 3(F-2)$. Subdominant classes are special configurations with lesser edges and vertices.

- All classes are connected by a finite sequence of 2-2 Pachner moves. The move (245) connects classes with the same number of edges, whereas the move (246) permits to descend to a subdominant class.
- The lowest-dimensional class corresponds to a maximal number of triangular faces, a condition which minimizes the number of vertices. When *all* the faces are triangular, the polytope can be seen as a collection of tetrahedra glued together, and with matching conditions imposed along all shared internal triangles.

8.3 Large F and the hexagonal dominance

The number of classes grows very fast with F (see for instance [81] for a tabulation). Thus far, we have been able to characterize the class looking at how many faces have a certain valency. The valency distribution is enough to identify the class at small F, however as we increase F we find classes with the same valency distribution, but which differ in the way the faces are connected. To distinguish the classes one needs to identify the complete combinatorial structure of the polytope. This information is captured by the *adjacency matrix*, which codes the connectivity of the faces of the polytope. Below in Section 9.3 we will show how this matrix can be explicitly built as a function of areas and normals, and give some explicit examples.

An interesting question concerns the average valency of a face, defined as $\langle p \rangle = 2E/F$. A simple estimate can be given using the fact that the boundary of any polytope is a cellular decomposition of the two-sphere, therefore by the Euler formula $F - E + V = 2$. For the dominant classes, which are dual to triangulations, the additional relation $2E = 3V$ holds, hence $E = 3(F-2)$ and we get $\langle p \rangle = 6(1-2/F)$. For large F, we expect the polytope to be dominated by *hexagonal faces*. This expectation is immediately confirmed by a simple numerical experiment. A polytope with large number of faces can be drawn using Wolfram's Mathematica, and $\langle p \rangle$ can be evaluated.The specimen in Figure 10, for instance, has $F = 100$,normals uniformly distributed on a sphere, all areas equals,and $\langle p \rangle \sim 5.88$. Notice also from the image that there are no triangular faces, consistently with the fact that they tend to minimize the number of vertices and are thus highly non-generic configurations.

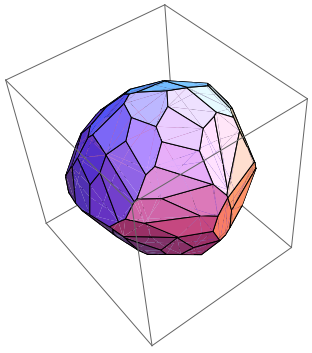


Figure 10: A generic polytope with $F = 100$. Most faces have valency 6. Triangles are nowhere to be seen.

9 Polytopes: reconstruction procedure and geometry

So far we have discussed how a point in $\mathcal{S}_F$ specifies a unique polytope, and the existence of different combinatorial structures. We now describe how the polytope can be explicitly reconstructed from areas and normals. The reconstruction will allow us to identify its class through the adjacency matrix, and evaluate completely its geometry, including the lengths of the edges and the volume.

Let us first give an overview of our strategy. From the definition (241), we know how to construct a polytope as the (internal region to the) intersection of F planes, identified by their normals n_i and translated by h_i from the origin. The contruction can be performed explicitly using an algorithm due to Lasserre [86], which permits to compute the lengths $\ell_{ij}(h,n)$ of all the edges of the polytope defined by h_i and n_i. Using this result, we derive a formula for the areas as $A_i(h,n)$. Inverting this formula, we get the heights $h_i(A,n)$ relative to a specific choice of areas and normals, and thus we can reconstruct the polytope in terms of the data (A,n).

It is worth adding that the problem of computing the volume of a given polyhedron is a complex and well studied topic in computational mathematics [111, 112], hence better procedures than the one used here could in principle be found. However, the usual starting point for common algorithms is the knowledge of the coordinates of vertices, or the system of inequalities (241). Therefore the methods need to be adapted to obtain formulas in terms of areas and normals. We found Lasserre's algorithm to be the most compatible with this necessity.[11]

We now review and adapt to 3d Lasserre's procedure.

9.1 Lasserre's algorithm

The basic idea of the reconstruction algorithm is to compute the length of an edge as the length of an interval in coordinates adapted to the edge. Consider the i-th face. From the defining inequalities (241), we know that points on this face satisfy

$$n_i \cdot x = h_i \tag{247a}$$

$$n_j \cdot x \leq h_j, \quad i \neq j. \tag{247b}$$

We introduce coordinates y_i adapted to the face, that is

$$n_i \cdot y_i = 0, \qquad y_i = x - (x \cdot n_i) n_i. \tag{248}$$

Using (247a) we get $x = h_i n_i + y_i$, which inserted in (247b) gives

$$y_i \cdot n_j \leq r_{ij}, \quad i \neq j \ , \tag{249}$$

where we have defined

$$r_{ij} \equiv h_j - (n_i \cdot n_j) h_i \ . \tag{250}$$

Hence, the i-th face can characterized either in terms of the x or the y_i coordinates,

$$\left\{ \begin{array}{l} x \cdot n_i = h_i \\ n_j \cdot x \leq h_j, \quad i \neq j \end{array} \right. \quad \longrightarrow \quad \left\{ \begin{array}{l} y_i \cdot n_i = 0 \\ y_i \cdot n_j \leq r_{ij}(h,n), \quad i \neq j \end{array} \right. \tag{251}$$

[11] Numerical algorithms for the volume and shape reconstruction from areas and normals are developed in the study of extended Gaussian images in informatics [113]. However there are no analytical results.

Notice that $r_{ij}/\sqrt{1-(n_i\cdot n_j)^2}$ is the distance of the edge ij from the projection of the origin on the i-th face.

The next step is to iterate this process and describe an edge in terms of its adapted coordinates. We start from the i-th face again, and assume that it is connected to the face j, so that the two faces share an edge. Points on the edge ij between the i-th and the j-th face satisfy

$$y_i\cdot n_i = 0 \tag{252}$$

$$y_i\cdot n_j = r_{ij} \tag{253}$$

$$y_i\cdot n_k \le r_{ik}, \qquad k\neq i,j. \tag{254}$$

As before, we introduce coordinates z_{ij}, adapted to the edge,

$$n_i\cdot z_{ij} = n_j\cdot z_{ij} = 0, \qquad z_{ij} = y_i - [n_j - (n_i\cdot n_j)n_i]\,\frac{y_i\cdot n_j}{1-(n_i\cdot n_j)^2}. \tag{255}$$

Using (253) we get that for a point in the edge

$$y_i = [n_j - (n_i\cdot n_j)n_i]\,\frac{h_j - h_i(n_i\cdot n_j)}{1-(n_i\cdot n_j)^2} + z_{ij}. \tag{256}$$

Plugging this in (254) gives

$$z_{ij}\cdot n_k \le b_{ij,k}, \tag{257}$$

where we have defined

$$b_{ij,k} \equiv h_k - (n_i\cdot n_k)h_i - \frac{(n_j\cdot n_k)-(n_i\cdot n_j)(n_i\cdot n_k)}{1-(n_i\cdot n_j)^2}\,[h_j - h_i(n_i\cdot n_j)]\ . \tag{258}$$

Summarizing as before, going to adapted coordinates the edge is defined by

$$\left\{\begin{array}{l} y_i\cdot n_i = 0 \\ y_i\cdot n_j = r_{ij}(h,n) \\ y_i\cdot n_k \le r_{ik}(h,n), \quad k\neq i,j. \end{array}\right. \longrightarrow \left\{\begin{array}{l} z_{ij}\cdot n_i = 0 \\ z_{ij}\cdot n_j = 0 \\ z_{ij}\cdot n_k \le b_{ij,k}(h,n), \quad i\neq j\neq k \end{array}\right. \tag{259}$$

At this point we are ready to evaluate the length of each edge. To that end, we parametrize the z_{ij} coordinate vector in terms of its norm, say λ, and its direction which is given by the wedge product of the two normals,

$$z_{ij} = \lambda\frac{n_i\wedge n_j}{\sqrt{1-(n_i\cdot n_j)^2}}. \tag{260}$$

If we define

$$a_{ij,k} \equiv \frac{n_i\wedge n_j\cdot n_k}{\sqrt{1-(n_i\cdot n_j)^2}}, \tag{261}$$

we can rewrite the inequalities in (259) as

$$\lambda a_{ij,k} \le b_{ij,k}. \tag{262}$$

Finally, the length of the edge is the length of the interval determined by the tightest set of inequalities, i.e.

$$\min_{k|a_{ij,k}>0}\left\{\frac{b_{ij,k}}{a_{ij,k}}\right\} - \max_{k|a_{ij,k}<0}\left\{\frac{b_{ij,k}}{a_{ij,k}}\right\}. \tag{263}$$

Here the minimum is taken over all the k's such that $a_{ij,k}$ is positive, and the maximum over all the k's

such that $a_{ij,k}$ is negative. This quantity is symmetric [86] and well-defined.[12]

Now comes a key fact: (263) can be defined for *any pair of faces* ij, not only if their intersection defines an edge in the boundary of the polytope. However, it is *negative* every time the edge does not belong to the polytope [86]. Thanks to this property, we can consistently define the edge lengths for any pair of faces ij as

$$\ell_{ij}(h,n) = \max_k \left\{ 0, \min_{k|a_{ij,k}>0} \left\{ \frac{b_{ij,k}}{a_{ij,k}} \right\} - \max_{k|a_{ij,k}<0} \left\{ \frac{b_{ij,k}}{a_{ij,k}} \right\} \right\}. \tag{264}$$

The result is a matrix whose entries are the edge lengths as a functions of the normals and the heights, linear in the heights, and zero if the intersection is outside the polytope.

This formula completes Lasserre's algorithm, and permits one to reconstruct the polytope from the set (h_i, n_i). To achieve a description in terms of areas at the place of heights, we use (264) to compute the areas of the faces. To do so, we use the projection of the origin on the face to divide it into triangles. Recall the Lasserre's procedure has provided us with the distance between an edge and the projected origin, see (251). We thus can write

$$A_i = \frac{1}{2} \sum_{\substack{j=1 \\ j\neq i}}^{F} \frac{r_{ij}}{\sqrt{1-(n_i \cdot n_j)^2}} \, \ell_{ij}. \tag{265}$$

Notice that both $r_{ij}(h,n)$ from (249) and $\ell_{ij}(h,n)$ from (264) are linear in the heights. Hence, the area is a quadratic function,

$$A_i(h,n) = \sum_{j,k=1}^{F} M_i^{jk}(n_1, \ldots, n_F) h_j h_k, \tag{266}$$

where M_i is a complicated matrix depending only on the normals. This quadratic system can be inverted to obtain $h_i(A,n)$. The existence of a solution with $h_i > 0 \ \forall i$ is guaranteed by Minkowski's theorem. However, the solution is not unique: in fact, we have the freedom of moving the origin around inside the polytope, thus changing the value of the heights without changing the shape of the polytope.

To our knowledge there are no systematical procedure that can be used to solve a generic quadratic system.In order to invert it, we canfix an origin in such a way that we can put as much h to zero as possible. One can be tempted to put the origin on a vertex or an edge but this is impossible to do without knowing a priori the combinatorial structure of the polytope. We can at most put the origin on a face and put to zero one h.

Finally, from the inverse we derive the lengths as functions of areas and normals, which with a slight abuse of notation we still denote in the same way,

$$\ell_{ij}(A,n) = \ell_{ij}(h(A,n),n). \tag{267}$$

These expressions are well-defined and can be computed explicitly.

9.2 Volume of a polytope

As discussed, a set of areas and normals that satisfy the closure condition fully determine the shape of a polytope. Let us call $\mathcal{P}(A_i, n_i)$ the convex subset of $\mathbb{R}^3$ corresponding to the polytope. Its volume is simply the integral of the Euclidean volume density on this region:

$$V(A_i, n_i) = \int_{\mathcal{P}(A_i,n_i)} d^3\vec{x} \ . \tag{268}$$

[12] It diverges if the normals are parallel, but this case is excluded since the two faces would not share an edge.

An interesting question is how to compute efficiently the volume integral (268). The simplest way is to use the algorithm described in the previous section: we chop the region $\mathcal{P}(A_i, n_i)$ into pyramids with a common vertex in its interior and bases given by the faces of the polytope. In this way the volume is just the sum of the volumes of the pyramids, i.e.

$$V(A_i, n_i) = \frac{1}{3} \sum_{i=1}^{F} h_i \, A_i \ . \tag{269}$$

Here $h_i = h_i(A, n)$ are the heights of the pyramids expressed in terms of the areas and normals via Lasserre's algorithm. Notice that since the volume is manifestly invariant under rotations, it can also be written directly as a function of the reduced variables (μ_k, θ_k). To do so, one only needs to use the relation $n_i = n_i(\mu_k, \theta_k)$, which is straightforward to derive once a reference frame is chosen.

The volume of the polytope as a function of areas and normals has a number of interesting properties that we now discuss:

C1. Clearly, as a function of areas and normals, the volume is either positive or vanishing. In particular, for given areas, it vanishes only when the normals n_i lie in a plane. To be precise, Minkowski theorem and the algorithm discussed in the previous section make sense only for non-coplanar normals. However, the limit of coplanar normals exists and the volume tends to zero in this limit[13].

C2. For fixed areas A_i, the volume is a bounded function of the normals. We call $V_{\max}(A_i)$ the volume of the polytope with maximum volume[14],

$$V_{\max}(A_i) \equiv \sup_{n_i} V(A_i, n_i) \ . \tag{270}$$

In particular, $V_{\max}(A_i)$ is smaller that the volume of the sphere that has the same surface area as the polytope. Therefore we have the bound

$$0 \leq V(A_i, n_i) < \frac{\left(\sum_i A_i\right)^{\frac{3}{2}}}{3\sqrt{4\pi}} \ . \tag{271}$$

C3. *Face-consistency.* Setting to zero the area of a face such that the result is still a non-degenerate polytope, (269) automatically measures the volume of the reduced polytope with $F - 1$ faces.

In conclusion, a point in $\mathcal{S}_F$ determines uniquely the whole geometry of a polytope and in particular its edge-lengths l_{ij} (264) and its volume (269). Now we show how these data can be used to identify the class of the polytope.

[13] In order to to see this we can exploit the property (C3) and notice that we can obtain a general F valent coplanar configuration from a $F+1$ pyramid configuration in the limit of zero base's area. Consider for example a cross shaped set of four normals (two by two collinear and opposite and one set orthogonal to the other). Using the above construction we take five normals that correspond to a square base pyramid and, if we call ϵ the area of the base, we perform the limit $\epsilon \to 0$ keeping constant the area of the triangular faces. To keep the area of a triangle fixed to A in the limit process the height has to be $h = \sqrt{\frac{A^2}{\epsilon} - \frac{\epsilon}{4}} \approx A/\sqrt{\epsilon}$. In this limit the configuration of the four remaining normals becomes planar (i) and the volume associated to this configuration vanishes $V = \frac{h\epsilon}{3} = \frac{\sqrt{\epsilon}}{3} \to 0$. One can be tempted to say that this result comes trivially from (269) with all the heights equals to zero. This could be tricky, in fact (269) is not well defined for degenerate configuration, but thanks to the above argument we can extend (269) also for degenerate configuration setting all h to zero.

[14] Notice that there can be more than one polytope that attain maximum volume. For instance, in the case $F = 4$, there are two parity related tetrahedra that have the maximum volume.

9.3 Adjacency matrix

The adjacency matrix A of the polytope is defined as

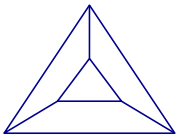
$$A_{ij} = \begin{cases} 1 & \text{if the faces } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases} \qquad i,j = 1,\ldots,F \tag{272}$$

Notice that A_{ij} coincides with the matrix ℓ_{ij} in (264) with all the non-zero entries normalized to 1: the reconstruction algorithm gives us the adjacency matrix for free.

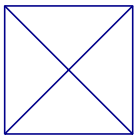
A_{ij} contains information on the connectivity of the faces as well as on the valence of each face, thus the class of the polytope can be identified uniquely from it. The valence p_i of the face i can be extracted taking the sum of the columns for each row,

$$p_i = \sum_{j=1}^{F} A_{ij}. \tag{273}$$

For example, for the two classes with $F = 5$ of Fig.7 we have



$$\longrightarrow \quad A = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix} \quad \longrightarrow \quad p = (3,3,4,4,4)$$



$$\longrightarrow \quad A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix} \quad \longrightarrow \quad p = (4,3,3,3,3)$$

From graph theory [114], we known that (272) has a number of interesting properties that can be related to the geometrical parameters of the polytope. For instance, the number of walks from the face i to the face j of length r is given by the matrix elements of the r-th power $(A^r)_{ij}$. From this property we deduce that the number E of edges of the polytope is

$$E = \frac{1}{2}\,\mathrm{Tr}A^2 = \frac{1}{2}\sum_i p_i. \tag{274}$$

This expression generalizes the value $E = 3(F-2)$ valid for the dominant classes.

Higher traces are related to the number of loops of a given lengh. For instance, the number of closed loops of length 3 is given by $(1/6)\,\mathrm{Tr}A^3$.

9.4 Mapping the phase space

Now that we have a tool to distinguish the classes, we can go back to our initial program of mapping the phase space, and do it in a quantitative way. This requires inverting (266), which we did numerically with the help of Wolfram's Mathematica. We subdivided the phase space into a regular grid, and had Mathematica computing the adjacency matrix of the area-normal configurations lying at the vertices of the grid. This associates a unique class to each cell of the phase space. This information can be colour-coded, and the result for a specific example is shown in Fig.11. Notice that in this way we only realize an

approximate mapping, where by construction we have measure-zero probability of hitting a subdominant class, thus the latter are absent in the mapping. Furthermore, although (266) is always invertible, the specific algorithm one uses might get in trouble on some specific configurations. Configurations for which our numerical algorithm failed are holes in the pictures. This limitation can be easily improved with a better algorithm inversions, or by choosing a configuration slightly off the vertex of the grid. However it is irrelevant for us to improve it at this stage. What we care to show in this paper is that it is possible to do the mapping in a quantitative way.

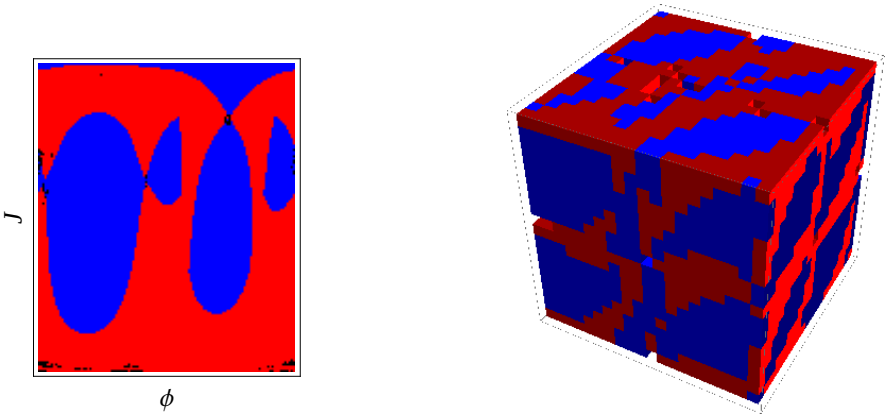


Figure 11: Mappings of subspaces of $\mathcal{S}_6$ at fixed external areas (respectively 9,10,11,12,13,13). Left panel: a pair (J, ϕ) (the other phase space variables are fixed to $J_1 = 15$, $\phi_1 = \frac{7}{10}\pi$, $J_2 = 13$, $\phi_2 = \frac{13}{10}\pi$). Right panel: the three angles ϕ_i for fixed J_i (respectively 15,13,17). The mappings have a finite resolution: each colored box corresponds to a single configuration. The cuboids are in blue, the pentagonal wedges in red. The holes are configurations for which our numerical algorithm to invert (266) failed.

9.5 Shape matching conditions

Knowing the complete geometry of the polytopes allows us also to address the following situation. Suppose we are given two polytopes in terms of their areas and normals, and suppose we want to glue them by a common face. Even if we choose the area of the common face to be the same, there is no guarantee that the shape of the face will match: The two sets of data will in general induce different shapes of the face. That is, the face has the same area but it can be two different polygons altogether. In order to glue the polytopes nicely, one needs shape matching conditions guaranteeing that the shared face has the same geometry in both polytopes.

If both polytopes are tetrahedra, the problem has been solved in [52]. One uses the fact that the shape of the common triangle will match if two lengths, or two internal angles, are the same (only, two, since the area is already matching). In particular, the internal angles α can be expressed in terms of the 3d dihedral angles of the tetrahedron as follows,

$$\cos \alpha^i_{jk} = \frac{\cos \phi_{ij} + \cos \phi_{ik} \cos \phi_{jk}}{\sin \phi_{ik} \sin \phi_{jk}}. \tag{275}$$

Here the faces i, j and k all share a vertex, and α^i_{jk} is the angle between the edge ij and the edge ik inside the triangle i. See left panel of Figure 4. Consider now the adjacent tetrahedron. Its geometry induces for the same angle the value

$$\cos \alpha^i_{j'k'} = \frac{\cos \phi'_{ij'} + \cos \phi'_{ik'} \cos \phi'_{j'k'}}{\sin \phi'_{ik'} \sin \phi'_{jk'}}. \tag{276}$$

Hence, for the shape to match it is sufficient to require

$$\mathcal{C}_{kl,ij}(\phi) \equiv \cos\alpha^i_{jk} - \cos\alpha^i_{j'k'} = 0 \tag{277}$$

for two of the three angles of the triangle.

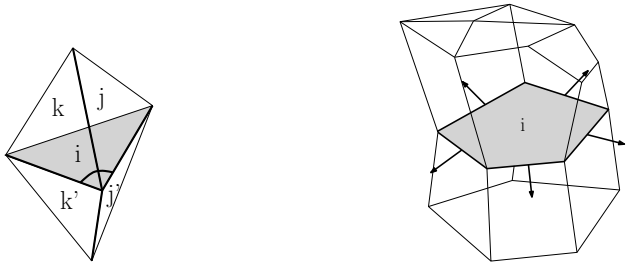


Figure 12: The geometric meaning of equation (277): the 2d angle $\alpha_{ij,kl}$ belonging to the shaded triangle can be expressed in terms of 3d angles associated the thick edges of the tetrahedron k, or equivalently of the tetrahedron l.

The simplicity of the conditions (277) is a consequence of the fact that two triangles with the same area are congruent if two angles match. For the general case, the face to glue is now a polygon and the number of conditions greater. One needs to make sure that the valency p of the polygon is the same. Then, the number of independent parameters of a polygon on the plane is $2p-3$, hence giving the edge lengths is not enough, and $p-2$ additional conditions are needed. A convenient procedure is then the following. Consider the face with the matching area, and one polytope. First of all, we know all the edge lengths ℓ_{ij} of the face viewed from the polytope. Then, for all the j such that $\ell_{ij} \neq 0$, we consider the face normals n_j projected on the plane of the i-th face,

$$\tilde{n}_j = \frac{n_j - (n_i \cdot n_j)n_i}{|n_j - (n_i \cdot n_j)n_i|} = \frac{n_j - \cos\phi_{ij} n_i}{\sin\phi_{ij}}. \tag{278}$$

The set $(\ell_j, \tilde{n}_j)$ defines a unique polygon in the plane identified by n_i, thanks to a two-dimensional version of Minkowski's theorem. Next, we do the same with the second polytope, obtaining a second set $(\ell'_j, \tilde{n}'_j)$ living in the plane identified by n'_i. Finally, the shape matching conditions consist of imposing the equivalence of these two flat polygons up to rotations in three-dimensional space.

10 Relation to loop quantum gravity

Thus far we have been discussing classical properties of polytopes. In the rest of the paper, we discuss the relevance of polytopes for loop quantum gravity.

The relation between polytopes and loop quantum gravity comes from the key result that:

- intertwiners are the building blocks of spin-network states, an orthonormal basis of the Hilbert space of loop quantum gravity, [12]
- intertwiners can be understood as the quantization of the phase space of polytopes with fixed areas, i.e. of the Kapovich and Millson's phase space $\mathcal{S}_F$ [82, 102, 83] (see also [80, 84]). Notice that as usual, constructing the quantum theory implies a quantization of some classical quantities, in this case the areas.

Therefore an intertwiner can be seen as the state of a *quantum polytope* and spin-network states as a collection of quantum polytopes with adjacency relations prescribed by the spin-network graph.

In this section we discuss the quantum polytope, semiclassical states and twisted geometries.

10.1 The quantum polytope

The intertwiner space $\mathcal{H}_F = \mathrm{Inv}\big[\otimes_{i=1}^F V^{(j_i)}\big]$ can be seen as the Hilbert space of a quantum polytope. In this sense, our construction provides a generalization to arbitrary valence of the seminal work of Barbieri [77] where the quantum tetrahedron was introduced (see also [78]). The correspondence between classical quantities and their quantization is the following: up to a dimensionful constant, the generators $\hat{J}_i$ of $SU(2)$ acting on each representation space $V^{(j_i)}$ are understood as the quantization of the vectors $A_i n_i$, i.e. of the normals to the faces of the polytope normalized to the area. In LQG the dimensionful constant is chosen to be the Immirzi parameter times Planck's area,

$$A_i n_i \qquad \longrightarrow \qquad \hat{E}_i = 8\pi\gamma G\hbar\, \hat{J}_i. \tag{279}$$

The closure condition (242) for the polytope becomes the the definition of intertwiner

$$\sum_i \hat{E} \ = 0 \ , \tag{280}$$

and corresponds to the Gauss constraint in classical General Relativity in Ashtekar-Barbero variables.

10.2 Coherent intertwiners and semiclassical polytopes

Coherent intertwiners for $\mathcal{H}_F$ were introduced in [100] and developed in [102, 103]. They are defined as

$$\|j_i, n_i\rangle = \int \mathrm{d}g\ D^{(j_1)}(g)|j_1, n_1\rangle \cdots D^{(j_F)}(g)|j_F, n_F\rangle, \tag{281}$$

where $|j, n\rangle$ is an SU(2) Perelomov coherent state. Here, the unit-vectors n_i are assumed to close, $\sum_i j_i n_i = 0$. These states can be shown to be an overcomplete basis of the Hilbert space $\mathcal{H}_F$, this being a consequence of Guillemin-Sternberg's theorem that quantization commutes with reduction, as shown in [102, 103].

As the states $\|j_i, n_i\rangle$ are invariant under rotation of the vectors n_i, they can be label by the Kapovich and Millson canonical variables (μ_k, θ_k): coherent intertwiners are labeled by a point in the phase space $\mathcal{S}_F$, i.e. by a classical polytope.

From the point of view of geometric quantization, a convenient parametrization of the phase space $\mathcal{S}_F$ is in terms of $F-3$ complex numbers Z_k [102]. They are defined as follows. Let us consider the stereographic projection of a unit-vector $n=(n_x,n_y,n_z)$ into the complex plane,

$$z=-\frac{n_x-in_y}{1-n_z}=-\tan\frac{\theta}{2}\,e^{-i\phi}, \tag{282}$$

where θ and ϕ are the zenith and azimuth angles of S^2, and we have chosen to project from the south pole. The $F-3$ complex variables Z_k are defined in terms of the stereographic projection z_i of the normals n_i as their cross-ratio

$$Z_k=\frac{(z_{k+3}-z_1)(z_2-z_3)}{(z_{k+3}-z_3)(z_2-z_1)}\quad,\quad k=1,\dots,F-3\,. \tag{283}$$

In the following we use Z_k as labels for the coherent intertwiners,

$$|j_i,Z_k\rangle=\|j_i,n_i(Z_k)\rangle\,. \tag{284}$$

The variables Z_k provide a complex polarization[15] of the phase space $\mathcal{S}_F$. The advantage of these variables is that coherent states can be obtained in an elegant way through geometric quantization [102]. The states $|j_i,Z_k\rangle$ are holomorphic in Z_k and the resolution of the identity is

$$\mathbb{1}=\int \mathrm{d}\mu(Z)\,|j_i,Z_k\rangle\langle j_i,Z_k|\,, \tag{285}$$

where the measure

$$\mathrm{d}\mu(Z)=8\pi^2\big(\prod_i\frac{2j_i+1}{2\pi}\big)\,K(Z_k,\bar{Z}_k)\prod_k\mathrm{d}^2Z_k \tag{286}$$

is defined by an integration kernel $K(Z_k,\bar{Z}_k)$ that depends parametrically on the spins j_i and is given explicitly in [102].

These states are "minimally" overcomplete, since they achieve a parametrization of the phase space in terms of reduced variables. The advantage is that each one of them now represents a coherent polytope (including degenerate cases), thus (285) is truly a decomposition of the intertwiner space in terms of polytopes. [16]

The states $|j_i,Z_k\rangle$, provide coherent states for the space of intertwiners only,[17] and should not be confused with coherent states for loop quantum gravity. However, they do play a role. We now describe how polytopes are relevant to the full theory.

[15] Notice that it is not of the form $\mu_k+i\theta_k$.

[16] Another use of the states $|j_i,Z_k\rangle$ is to define an holomorphic representation of the quantum algebra on functions $\psi(Z_k)\equiv\langle j_i,\bar{Z}_k|\psi\rangle$, see [83]. We will not use this representation in this paper.

[17] Recently [115, 116, 117] attention has been given to a second space for which polytopes are relevant. This is a sum of intertwiner spaces such that the total spin is fixed,

$$\mathcal{H}_J=\bigoplus_{\substack{j_1..j_F\\ \sum_i j_i=J}}\mathrm{Inv}\big[\otimes_{i=1}^F V^{(j_i)}\big]. \tag{287}$$

The interest in this space is that it is a representation of the unitary group U(F). Vectors in this space represent quantum polytopes with fixed number of faces and fixed total area, but fuzzy individual areas as well as shapes as before. Coherent states for (287) can be built using U(F) coherent states [116]. Like the LS states, these are also peaked on classical polytopes. The study in this paper is thus relevant also for those.

10.3 Coherent states on a fixed graph and twisted geometries

To relate polytopes to loop quantum gravity, consider a truncation of the theory to a single graph Γ, with L links and N nodes. The associated gauge-invariant Hilbert space is $\mathcal{H}_\Gamma = L_2[SU(2)^L/SU(2)^N]$, which decomposes in terms of intertwiner spaces $\mathcal{H}_n \equiv \mathrm{Inv}[\otimes_{l\in n} V^{(j_l)}]$ as

$$\mathcal{H}_\Gamma = \oplus_{j_l} \left(\otimes_n \mathcal{H}_n\right). \tag{288}$$

This Hilbert space is the quantization of a classical space $S_\Gamma = T^*\mathrm{SU}(2)^L//\mathrm{SU}(2)^N$, which corresponds to (gauge-invariant) holonomies and fluxes associated with links and dual faces of the graph. The important result which is relevant for us is that this space, a symplectic manifold up to singular points [94], admits a decomposition analogue to (288). In fact, it can be parametrized as [79]

$$S_\Gamma = \bigtimes_l T^*S^1 \bigtimes_n \mathcal{S}_{F(n)}, \tag{289}$$

where T^*S^1 is the cotangent bundle to a circle, F the valency of the node n, and $\mathcal{S}_F$ is the space of shapes of Kapovich and Millson introduced earlier.

The parametrization is achieved through an isomorphism between holonomy-fluxes and a set of variables dubbed "twisted geometries". These are the assignment of an area A_l and an angle ξ_l to each link, and of F normals n_l, satisfying the closure condition (242), to each node. See [79, 53] for details and discussions.

As we have seen in the previous Sections, each configuration of (non-degenerate) closed normals defines a unique polytope with fixed external areas A_l. Hence, the twisted geometry parametrization of the phase space allows us to bridge from holonomy-fluxes to polytopes around each node. The reconstruction can be done explicitly first mapping holonomies and fluxes to the variables of twisted geometries, and then using areas and normals to reconstruct the polytopes as described in section 9. The extra angles ξ_l carry information on the extrinsic geometry, needed to define the parallel transport between adjacent polytopes.

The isomorphism (289) and the unique correspondence between closed normals and polytopes means that *each classical (non-degenerate) holonomy-flux configuration on a fixed graph can be visualized as a collection of adjacent polytopes, each one with its own frame, and with a notion of parallel transport between any two polytopes.* Just as the intertwiners are the building blocks of the quantum geometry of spin networks, polytopes are the building blocks of the classical phase space (289) in the twisted geometries parametrization.

Now, what is the relevance of this geometric construction to loop quantum gravity? Coherent states for loop quantum gravity have been introduced and extensively studied by Thiemann and collaborators [93, 94, 95]. Although the states for the full theory have components on each graph, it is often convenient to truncate the theory on a fixed graph. This truncation provides a useful computational tool, to be compared to a perturbative expansion, and has found many applications, from the study of propagators and n-point functions [128] to cosmology [127]. In many of these applications, control of the semiclassical limit requires a notion of semiclassical states in the truncated space $\mathcal{H}_\Gamma$.

The truncation can only capture a finite number of degrees of freedom, thus such semiclassical states can not be peaked on continuous, smooth geometries. On the other hand, it is useful to visualize them as peaked on some discrete geometries, approximation of a smooth one on a cellular decomposition dual to Γ. Twisted geometries implement this picture, therefore a semiclassical state can be seen as a collection of polytopes glued together.[18]

[18] This interpretation is independent of the details of the semiclassical states. All that is required is that they are properly

There are two subtleties with this geometric picture that should be kept in mind. The first one is that the polyhedral picture is only valid for non-degenerate configurations, i.e. when the normals are not coplanar. Degenerate configurations correspond instead to polygons living on the plane shared by the normals. Holonomy-flux configurations leading to coplanar normals in one or more node describe a discrete 3d geometry with degenerate regions.

The second subtlety is the fact that twisted geometries have a peculiar characteristic which justifies their name: they define a metric which is locally flat, but *discontinuous*. To understand this point, consider the link shared by two nodes. Its dual face has area proportional to A_l. However, the shape of the face is determined *independently* by the data around each node (i.e. the normals and the other areas), thus generic configurations will give two different shapes. In other words, the reconstruction of two polytopes from holonomies and fluxes does not guarantee that the shapes of shared faces match. Hence, the metric of twisted geometries is discontinuous across the face. See left panel of Figure 12.

One can also consider a special set of configurations for which the shapes match, see right panel of Figure 12. This is a subset of S_Γ where the shape matching conditions, discussed earlier in Section 9.5, hold. This subset corresponds to piecewise flat and continuous metrics. For the special case in which all the polytopes are tetrahedra, this is the set-up of Regge calculus, and those holonomies and fluxes describe a 3d Regge geometry: the reduced twisted geometries amount to the edge lengths and 3d dihedral angles. The relation between twisted geometry and Regge calculus implies that holonomies and fluxes carry *more* information than the phase space of Regge calculus. This is not in contradiction with the fact that the Regge variables and the LQG variables on a fixed graph both provide a truncation of general relativity: simply, they define two distinct truncations of the full theory. See [53] for a discussion of these aspects.

For an arbitrary graph, the shape-matching subset describes a generalization of Regge calculus on arbitrary cellular decompositions. In this case however the variables are not equivalent any longer to edge lengths, since as already discussed these do not specify uniquely the geometry of polytopes. Rather, such cellular Regge calculus uses areas and normals as fundamental variables. This is the set up of area-angle Regge calculus [106]. For the action, one can use the one of [106] adapted to the cellular decomposition, and with the gluing conditions generalized as discussed here.

peaked on a point in phase space. For instance, the twisted geometries parametrization has been applied to Thiemann's coherent states in [96]. Alternative coherent states based on a different complex structure suggested by (289) appear in [97]. These two papers also show how the coherent intertwiners enter the construction of complete coherent states.

11 On the volume operator

At the classical level, the volume of a polytope is a well-defined quantity. In this section we investigate the quantization of this quantity and its relation with the volume operators used in loop quantum gravity.

11.1 The volume of a quantum polytope

Let us consider the phase space $\mathcal{S}_F$ of polytopes with F faces of given area. The volume of the polytope is a well-defined function on this phase space, as discussed in Section 9.2. Coherent intertwiners provide a natural tool to promote this quantity to an operator.

In the following we use the parametrization of the phase space $\mathcal{S}_F$ in terms of the cross ratios Z_k. In particular, the F normals n_i are understood as functions of the cross-ratios, $n_i(Z_k)$. Accordingly we call $V(j_i, Z_k)$ the volume of a polytope with faces of area $A_i(j_i)$ and normals $n_i(Z_k)$,

$$V(j_i, Z_k) \equiv V(A(j_i), n(Z_k)) \ . \tag{290}$$

Let us consider now the Hilbert space of intertwiners $\mathcal{H}_F$ associated to the phase space $\mathcal{S}_F$. The volume of a quantum polytope can be defined in terms of coherent intertwiners $|j_i, Z_k\rangle$ and of the classical volume as follows:

$$\hat{V} = \int d\mu(Z_k)\ V(j_i, Z_k)\ |j_i, Z_k\rangle\langle j_i, Z_k| \ . \tag{291}$$

This integral representation of the operator in terms of its classical version[19] is of the kind considered originally by Glauber [122] and Sudarshan [123]. It has a number of interesting properties that we now discuss:

Q1. The operator $\hat{V}$ is positive semi-definite, i.e.

$$\langle\psi|\hat{V}|\psi\rangle = \int d\mu(Z_k)\ V(j_i, Z_k)\ |\langle j_i, Z_k|\psi\rangle|^2 \ \ \geq 0 \ , \tag{292}$$

for every $|\psi\rangle$ in $\mathcal{H}$. This is a straightforward consequence of the fact that the classical volume is a positive function, $V(j_i, Z_k) \geq 0$.

Q2. $\hat{V}$ is a bounded operator in $\mathcal{H}_F$. Its norm $||\hat{V}|| = \sup_\psi \langle\psi|\hat{V}|\psi\rangle/\langle\psi|\psi\rangle$ is bounded from above by the maximum value of the classical volume of a polytope with fixed areas,

$$\langle\psi|\hat{V}|\psi\rangle = \int d\mu(Z_k)\ V(j_i, Z_k)\ |\langle j_i, Z_k|\psi\rangle|^2 \ \ \leq \ \ \sup_{Z_k} V(j_i, Z_k) \ \equiv V_{max}(j_i) \ . \tag{293}$$

Q3. *Spin j-consistency.* Let us consider the operator $\hat{V}$ defined on the Hilbert space $\mathcal{H}_{F+1}$ associated to spins $j_1, \dots, j_F, j_{F+1}$, and the one defined on the Hilbert space $\mathcal{H}_F$ associated to spins $j_1, \dots, j_F$. When the spin j_{F+1} vanishes, the two operators coincide. This is a consequence of the fact that the classical volume of a polytope with $F+1$ faces coincides with the volume of a polytope with F faces when one of the areas is sent to zero.

[19] In the literature [129], the classical function $V(j_i, Z_k)$ is called the P-symbol of the operator $\hat{V}$. On the other hand, the expectation value of the operator $\hat{V}$ on a set of coherent states, i.e.

$$Q_V(j_i, Z_k) \equiv \langle j_i, Z_k|\hat{V}|j_i, Z_k\rangle \ ,$$

is called the Q-symbol. When the P-symbol and the Q-symbol of an operator exist, then the operator is fully determined by either of them. The properties of these symbols and of the operator they define have been studied by Berezin in [130],[131]

These three properties are the quantum version of C1, C2, C3 discussed in section 9.2. Moreover, using the fact that for large spins the overlap of two coherent intertwiners tend to a delta-function,

$$|\langle j_i, Z_k | j_i, Z'_k\rangle|^2 \to \delta(Z_k, Z'_k) \ , \tag{294}$$

we have that the expectation value $\langle \hat{V}\rangle_{Z_k}$ of the volume operator on a coherent state $|j_i, Z_k\rangle$ reproduces the volume of the classical polytope with shape (j_i, Z_k),

$$\langle \hat{V}\rangle_{Z_k} \equiv \frac{\langle j_i, Z_k|\hat{V}|j_i, Z_k\rangle}{\langle j_i, Z_k|j_i, Z_k\rangle} \approx V(A_i(j_i), n_i(Z_k)) \ . \tag{295}$$

This fact allows to estimate the largest eigenvalue of the volume: in the large spin limit, the largest eigenvalue is given by $V_{max}(A_i)$, the volume of the largest polytope in $\mathcal{S}_F$.

The spectrum of the operator $\hat{V}$ that can be computed numerically. Let us focus on the case $F = 4$ for concreteness. The matrix elements of $\hat{V}$ in the conventional recoupling basis are given by[20]

$$V_{kk'} = \langle j_i, k|\hat{V}|j_i, k'\rangle = \int \mathrm{d}\mu(Z)\, V(j_i, Z)\, \langle j_i, k|j_i, Z\rangle\langle j_i, Z|j_i, k'\rangle, \tag{296}$$

The matrix $V_{kk'}$ can be diagonalized numerically to obtain its eigenvalues. We focused for simplicity on the simplest case where all the four spins j_i are equal. The results using Wolfram's Mathematica are shown in Fig.13 and confirm that the maximum eigenvalue is below the volume of the regular tetrahedron.

The spectrum has a gap. One of the interesting questions to investigate in the future is whether this gap survives at higher valency, or it is dumped as for the standard volume operator [121].

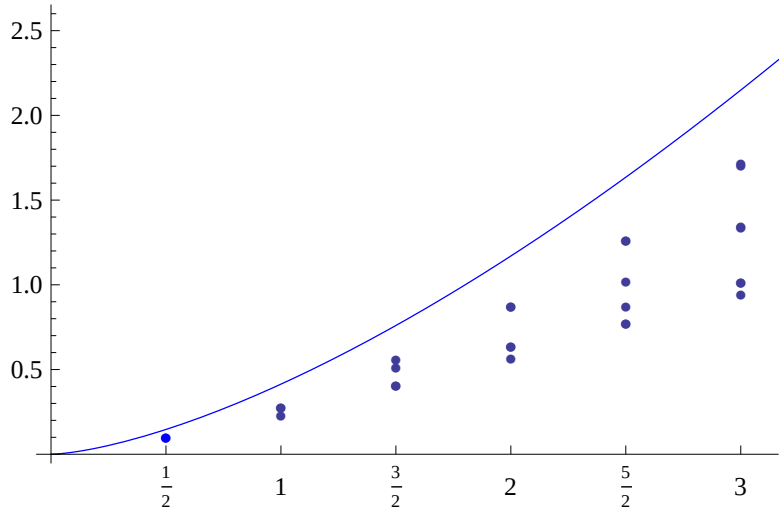


Figure 13: The spectrum of $\hat{V}$, blue circles, compared with Barbieri and the classical equilateral volume. The spectra are bounded from above by the classical maximal volume, that is the equilateral configuration. The red dashed line is the equilateral volume with $A = \sqrt{j(j+1)}$, to which Barbieri's asymptotes, whereas the blue continuous line is the equilateral volume with $A = j$, to which $\hat{V}$ asymptotes.

It is interesting to notice that the volume operator introduced above commutes with the *parity*

[20] The overlap $\langle j, k|j_i, Z\rangle = \frac{(-1)^{2k}}{2j+k+1}\frac{(2j!)^2}{(2j+k)!(2j-k)!}L_k(1-2Z)$, where L_k is the k-th Legendre polynomial, was computed in [102].

operator. This is the operator that sends the normals to their opposite,

$$\hat{\mathcal{P}}|j,n\rangle = |j,-n\rangle. \tag{297}$$

In terms of the stereographic projection, the maps $n \mapsto -n$ amounts to $z \mapsto -1/\bar{z}$, thus its action on coherent intertwiners labeled by the single cross-ratios Z is simply

$$\hat{\mathcal{P}}|j_i, Z\rangle = |j_i, \bar{Z}\rangle \ . \tag{298}$$

Notice that $V(j_i, Z_k) = V(j_i, \bar{Z}_k)$ thanks to the invariance of the classical volume under parity. Moreover the measure $\mathrm{d}\mu(Z_k)$ is invariant under the transformation $Z_k \to \bar{Z}_k$. As a result, the operator (291) commutes with parity since

$$\hat{\mathcal{P}}\hat{V}\hat{\mathcal{P}}^\dagger = \int \mathrm{d}\mu(Z)\, V(j_i, Z)|j_i, \bar{Z}\rangle\langle j_i, \bar{Z}| = \int \mathrm{d}\mu(\bar{Z})\, V(j_i, \bar{Z})|j_i, Z\rangle\langle j_i, Z| = \hat{V} \ . \tag{299}$$

Clearly, there are other possible proposals for the volume of a quantum polytope. All of them share the same classical limit, but can have a different spectrum for small eigenvalues. An interesting variant to consider is the operator obtained as the square-root of the modulus of the oriented-volume square $\hat{U}$ defined below

$$\hat{U} = \int d\mu(Z_k)\ s(Z_k) V^2(j_i, Z_k)\ |j_i, Z_k\rangle\langle j_i, Z_k| \ , \tag{300}$$

where $s(Z_k)$ is the parity of the polytope, i.e. $s(Z_k) = \pm 1$ and $s(\bar{Z}_k) = -s(Z_k)$. The operator $\hat{U}$ now anticommutes with the parity therefore, under the assumption that its spectrum is non-degenerate, we have that its eigenvalues appear in pairs $\pm u$. In particular, when the Hilbert space $\mathcal{H}_F$ is odd-dimensional, a zero eigenvalue will be present. As a result, the spectrum of this alternate volume operator $\hat{V} = \sqrt{|\hat{U}|}$ has non-vanishing eigenvalues that are twice degenerate and eventually a non-degenerate zero eigenvalue for in odd dimension. In Figure 13 , its spectrum is shown and compared to the spectrum of the volume of a quantum tetrahedron introduced by Barbieri and defined as $\sqrt{|J_1 \cdot (J_2 \times J_3)|}$.

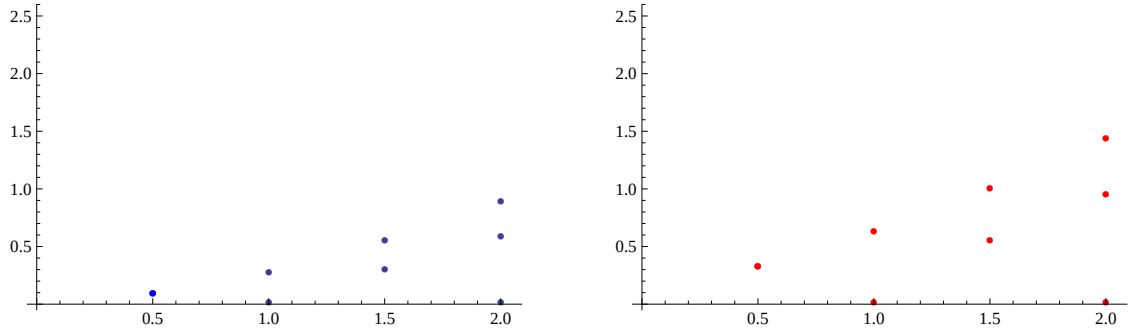


Figure 14: The spectrum of $\sqrt{|\hat{U}|}$, blue circles, compared with Barbieri and the classical equilateral volume. The spectra are bounded from above by the classical maximal volume, that is the equilateral configuration. The red dashed line is the equilateral volume with $A = \sqrt{j(j+1)}$, to which Barbieri's asymptotes, whereas the blue continuous line is the equilateral volume with $A = j$, to which $\hat{V}$ asymptotes.

11.2 LQG volume operator(s) and the quantum polytope

Clearly, this operators is not simply related to cubic flux operators, as the traditional Rovelli-Smolin and Ashtekar-Lewandowski ones. Furthermore, we do not know how to define them in the full theory,

that is prior to the introduction of a graph. On the other hand, operators of this type could play a role while studying approximations of the theory truncated on a given graph. Both operators are well-defined, thanks to the boundedness of $V(j, Z)$. We also remark that if we set one of the spins to zero, we automatically get the corresponding volume for valence $F-1$, a property inherited from the classical volume function. In this sense, it can be considered "cylindrical consistent".

The metric volume of a region R is a well-defined operator in loop gravity, corresponding to a quantization of the classical formula $V(R) = \int_R \sqrt{\det E}$. It has been studied extensively (e.g. [120, 121]), particularly because it enters Thiemann's construction of the Hamiltonian constraint [43] and thus it is relevant to understand the quantum dynamics of the theory. In particular, it was shown in [95] that the volume operator fails to have the right semiclassical limit on a fixed graph, unless one uses special coherent states defined only on 6-valent graphs. We would like here to review this result in the light of the polytopical perspective built so far.

To construct the volume operator one needs to rewrite the classical formula, which involves a product of triads, in terms of fluxes, which are the fundamental operators of the theory. The details of this procedure are not relevant here; let us just recall that two proposals have appeared in the literature, the original one by Rovelli and Smolin [11], and a more recent one by Ashtekar and Lewandowski [124], and that the second one is preferred by the "triad test" proposed in [92]. What is important for us is that in both cases one obtains a specific function of the fluxes, which is of the type of a square root of a triple product of fluxes. Then, recall that on coherent states each flux is peaked[21] on the vector $A_l n_l$ associated to a given link. In this way, the expectation value of the volume operator becomes a specific function of areas and dihedral angles among the normals. To cut a long story short, in [95] it was realized that this function gives the classical volume only when the states represent a parallelepipedoid.

This result should be now clearer in the light of the geometry of polytopes studied here. In fact, we have seen already that the volume of a polytope is a much more complicate function of areas and normals, than the triple product function appearing in [95]. If we consider the (unique) polytope defined by the given areas and normals, the triple product of $A_l n_l$ is proportional to its volume only in the case of a tetrahedron, and not for a generic number of faces. The only exception is the case of a parallelepipedoid, since this can be chopped in six tetrahedra of the same volume. Hence, no surprise that it only works for parallelepipedoids as claimed in [95].[22]

What we have learned on the polytopes shows that it makes no sense to expect that an operator is peaked on the volume of the polytope, because the type of polytope depends not only on the valency of the node, but on the *data themselves*. Therefore, the volume operator of loop quantum gravity can not be thought of as the volume of a quantum polytope. This is not a limitation of the operator per se: in fact, adding the interpretation of a flat polytope is artificial, and there is no reason to enforce it, especially since the theory is a continuum one.

On the other hand, the polytopical picture is interesting when one works with the truncation of the theory, and wants a discrete geometry at disposal to use as an approximation. In that case, one might be interested in a volume operator quantizing the classical formula for the polytope. This question can be addressed exploiting the knowledge of the classical expression (269), and the fact that a mapping between classical functions and quantum operators can be devised through the help of coherent states.

[21] The details of the peakedness depends on the specific coherent states chosen. However here we are interested only in the expectation values and not on the variance, therefore we discard discussing these types of difference.

[22] They could have also claimed that it works for a tetrahedron. However this would require a different normalization of the operator, normalization that the authors fix according to the "triad test" of [92].

12 On dynamics and spin foams

Spin foam models for the dynamics of loop quantum gravity are usually built starting from a discretization of the spacetime manifold in terms of a simplicial triangulation Δ. The study of semiclassical dynamics can benefit from the interpretation of LQG variables in terms of areas and normals, because it allows a direct relation to discrete general relativity on Δ. Specifically, it allows to compare the algebraic transition amplitudes to exponentials of the Regge action [132, 128, 133, 103]. On the other hand, complete transition amplitudes for LQG require the use of more general 2-complexes than those those dual to simplicial manifolds. A direct construction of the path integral for arbitrary graph has not been attempted so far. However Kaminski, Kisielowski and Lewandowski [125] have proposed an extension of the EPR model to arbitrary boundary graphs, based on a direct extension of the algebraic structure of the EPR model.

Just as Regge calculus is useful to study the semiclassical behavior on simplicial manifolds, a generalization to arbitrary cellular decompositions could be relevant to the theory, and allow us to test whether models such as the KKL one can be related to it.At the end of Section 10.3 we discussed cellular Regge calculus in 3d. More in general, the advantage in using n-simplexes in n dimensions is that they are "rigid", namely assigning their $n(n+1)/2$ edge lenghts fixes a unique flat metric. n-simplexes are polytopes with a minimal number of faces, $F = n+1$. On the other hand, generic polytopes require additional information, that is $nF - n(n+1)/2$. Hence, if one wants to formulate Regge calculus on polytopes, lengths can not be used as fundamental variables anymore. In 3d, they have to be replaced by areas and angles. These are precisely the twisted geometries, therefore the boundary data of a generic spin foam can still be interpreted in terms of twisted geometries, that is a collection of polytopes.

More subtle is the situation in the 4d bulk. A lesson from the recent asymptotics studies of the EPR models is that the amplitude is dominated by saddle points. When the boundary data are compatible with the existence of a (unique) flat 4-simplex, then the amplitude is effectively weighted the Regge action for the 4-simplex.

An extension of this result can be envisaged to more general graphs (e.g. for graphs with nodes of valence greater than four). Consider a graph that is dual to the boundary of a 4-polytope. Then a possibility would be that again the amplitude admits interesting saddle points at geometric configurations. These are, this time, configurations for which generic 3-polytopes glue to form a unique flat 4-polytope. We know the conditions to be satisfied: from Minkowski's theorem, volume and closed 4-normals. This translates into the fact that saddle points should exist (at least) when the boundary data defines uniquely these data. Similarly to what happens in the 4-simplex, that saddle points exist when 4-simplex exists.

Then, the amplitude could be related to a form of the Regge action specialized to the 4-polytope. Such action can be seen as an "effective" Regge action in which the internal edge lengths have been integrated out using the field equations.

13 Conclusion

The Hilbert space of loop quantum gravity is the direct sum of spaces associated to graphs Γ. A convenient approximation is to restrict only to one of these spaces that, in turn, is decomposed in intertwiner spaces. In this thesis we clarified details on the relation between polytopes and intertwiners, and concluded that an intertwiner can be seen unambiguously as the state of a *quantum polytope*.

We investigated the following applications of the polytopical correspondence:

- We gave a geometrical interpretation of coherent states: they are peaked on classical polytopes. The truncation of the theory only capture a finite number of degrees of freedom, thus semiclassical states can not be peaked on continuous, smooth geometries. On the other hand, we can visualize them as peaked on some discrete geometries, approximation of a smooth one on a cellular decomposition dual to Γ. Therefore a semiclassical state can be seen as a collection of polytopes glued together. This geometrical interpretation will be essential in the future development of the theory. In fact coherent states are the main tool in the semiclassical analysis and visualize them in a geometrical picture is a great advantage.

- We presented a new volume operator that in the large spin limit reproduce the volume of a classical polytope. This operator is also "cylindrical consistent", unlike the standard volume operators: if we set one of the spins to zero, we automatically get the corresponding volume for one less valence, a property inherited from the classical volume function. The technique used in the definition of our volume operator can also be used to define the quantum counterpart of other geometrical function on the space of polytopes. Moreover operators of this type could play a role while studying approximations of the theory truncated on a given graph.

- We discussed how to apply the polytopical interpretation to the study of the semiclassical limit of LQG, in particular commenting a connection between the quantum dynamics and a generalization of Regge calculus on polytopes. We proposed to study the large spin limit of an extension of EPR spin foam model. We expect that, under some condition, the amplitude is related to some "effective" Regge action where the cells are our polytopes. In addition this interpretation will play a key role, using a polytopical cellation of the spacetime manifold, in the derivation of the generalized spin foam model.

In the end, we studied the relation between LQG and quantum polytope, providing important tools that we hope will bring, in the future, new and useful insights in loop quantum gravity.

References

[1] M. H. Goroff and A. Sagnotti, Quantum Gravity At Two Loops, Phys. Lett. B160 (1985) 81.

[2] K. Eppley and E. Hannah, Found. Phys. 7, 51 (1977).

[3] P. M. Pearle and E. Squires, Gravity, Energy Conservation And Parameter Values In Collapse Models, arXiv:quant-ph/9503019.

[4] J. B. Hartle and S. W. Hawking, Phys. Rev. D28, 2960 (1983).

[5] For an introduction, see MB Green, JH Schwarz, E Witten: *Superstring theory* (Cambridge University Press, New York 1987). For up to date references, I refer to Gary Gibbon's plenary talk on string theory at this conference.

[6] J. M. Maldacena, The large N limit of superconformal field theories and supergravity, Adv. Theor. Math. Phys. **2** (1998) 231
E. Witten, Anti-de Sitter space and holography, Adv. Theor. Math. Phys. **2** (1998) 253
For a Review: O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri and Y. Oz, Large N field theories, string theory and gravity, Phys. Rept. **323** (2000) 183

[7] A Strominger, G Vafa, Microscopic Origin of the Bekenstein-Hawking Entropy, Phys Lett B379 (1996) 99–104.

[8] K Krasnov, "Geometrical entropy from loop quantum gravity", Phys Rev D55 (1997) 3505; "On statistical mechanics of Schwarzschild black hole", Gen Rel and Grav in print (1997).

[9] C Rovelli, "Black Hole Entropy from Loop Quantum Gravity", Phys Rev Lett 14 (1996) 3288–3291; "Loop Quantum Gravity and Black hole Physics", Helv Phys Acta 69 (1996) 582–611.

[10] A Ashtekar, J Baez, A Corichi, K Krasnov, Phys Rev Lett 80 (1998) 904-907.

[11] C Rovelli, L Smolin, "Discreteness of area and volume in quantum gravity", Nucl Phys B442 (1995) 593–622. Erratum: Nucl Phys B456 (1995) 734.

[12] C. Rovelli and L. Smolin, Loop Space Representation of Quantum General Relativity, Nucl. Phys. B **331** (1990) 80.

[13] J. Donoghue, Leading quantum correction to the Newtonian potential, Phys. Rev. D 50, 3874 (1994).

[14] J. F. Donoghue, Introduction to the Effective Field Theory Description of Gravity, arXiv:gr-qc/9512024.

[15] B. Hatfield, Quantum Field Theory of Point Particles and Strings, Perseus Publishin, 1998

[16] C. W. Misner, K. S. Thorne, J. A. Wheeler, Gravitation, San Francisco: W. H. Freeman, 1973

[17] R. Geroch, Domain of dependence, J. Math. Phys. 11, 437-449 (1970)

[18] A. N. Bernal and M. Sánchez, On smooth Cauchy hypersurfaces and Geroch's splitting theorem, Comm. Math. Phys. 243 (2003), no. 3, 461-470.

[19] S. W. Hawking and G. T. Horowitz, The Gravitational Hamiltonian, Action, Entropy and Surface Terms, Class. Quant. Grav. 13, 1487 (1996).

[20] S. W. Hawking and C. J. Hunter, The gravitational Hamiltonian in the presence of non-orthogonal boundaries, 1996 Class. Quantum Grav. 13 2735

[21] C. J. Isham, Quantum gravity and the problem of time, presented at Recent Problems in Mathematical Physics, Proceedings of the NATO Advanced Study Institute, Salamanca, 1992

[22] C.Rovelli, Quantum Gravity, Cambridge University Press, Cambridge, UK, 2004

[23] A. W. Wipf, Hamilton's Formalism For Systems With Constraints, arXiv:hep-th/9312078.

[24] T. Thiemann, Modern Canonical Quantum General Relativity, Cambridge Monographs on Mathematical Physics, Cambridge, UK, 2007

[25] M. Henneaux, C. Teitelboim, Quantization of Gauge Systems, Princeton University Press, Princeton, 1992

[26] C. J. Isham, in *Quantum Gravity 2, A Second Oxford Symposium* (eds Isham, C. J., Penrose, R. and Sciama, D. W.), Clarendon Press, Oxford, 1981.

[27] C. J. Isham, Modern Differential Geometry for Physicists, World Scientific Lecture Notes in Physics, 1999

[28] S. Holst, Barbero's Hamiltonian derived from a generalized Hilbert-Palatini action, Phys. Rev. D **53** (1996) 5966 [arXiv:gr-qc/9511026].

[29] A. Perez and C. Rovelli, Physical effects of the Immirzi parameter, Phys. Rev. D **73** (2006) 044013 [arXiv:gr-qc/0505081].

[30] L. Freidel, D. Minic and T. Takeuchi, Quantum Gravity, Torsion, Parity Violation and all that, Phys. Rev. D **72** (2005) 104002 [arXiv:hep-th/0507253].

[31] N. Barros e Sa, "Hamiltonian analysis of general relativity with the Immirzi parameter," Int. J. Mod. Phys. D **10** (2001) 261 [arXiv:gr-qc/0006013].

[32] J. Samuel, Is Barbero's Hamiltonian formulation a gauge theory of Lorentzian gravity?, Class. Quant. Grav. **17** (2000) L141 [arXiv:gr-qc/0005095].

[33] J. Lewandowski, A. Okolow, H. Sahlmann and T. Thiemann, Uniqueness of diffeomorphism invariant states on holonomy-flux algebras, Commun. Math. Phys. **267** (2006) 703 [arXiv:gr-qc/0504147].
C. Fleischhack, "Representations of the Weyl Algebra in Quantum Geometry," Commun. Math. Phys. **285** (2009) 67 [arXiv:math-ph/0407006].

[34] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourao and T. Thiemann, Quantization of diffeomorphism invariant theories of connections with local degrees of freedom, J. Math. Phys. **36** (1995) 6456 [arXiv:gr-qc/9504018].

[35] A. Ashtekar and C. J. Isham, Representations of the holonomy algebras of gravity and nonAbelian gauge theories, Class. Quant. Grav. **9** (1992) 1433 [arXiv:hep-th/9202053].
A. Ashtekar, J. Lewandowski, Representation theory of analytic holonomy C^* algebras, in: J.C. Baez (Ed.), Knots and Quantum Gravity, Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 1994.
A. Ashtekar and J. Lewandowski, Projective techniques and functional integration for gauge theories. J. Math. Phys. 36 (1995), pp. 2170-2191

[36] A. Ashtekar and J. Lewandowski, Quantum theory of geometry. I: Area operators, Class. Quant. Grav. **14** (1997) A55 [arXiv:gr-qc/9602046].

[37] A. Ashtekar and J. Lewandowski, Quantum theory of geometry. II: Volume operators, Adv. Theor. Math. Phys. **1** (1998) 388 [arXiv:gr-qc/9711031].

[38] A. Ashtekar, A. Corichi and J. A. Zapata,Quantum theory of geometry. III: Non-commutativity of Riemannian structures, Class. Quant. Grav. **15** (1998) 2955 [arXiv:gr-qc/9806041].

[39] R. De Pietri and C. Rovelli, Geometry Eigenvalues and Scalar Product from Recoupling Theory in Loop Quantum Gravity, Phys. Rev. D **54** (1996) 2664 [arXiv:gr-qc/9602023].

[40] A. Ashtekar and J. Lewandowski, Background independent quantum gravity: A status report, Class. Quant. Grav. **21** (2004) R53 [arXiv:gr-qc/0404018].

[41] A. Ashtekar, Singularity Resolution in Loop Quantum Cosmology: A Brief Overview, J. Phys. Conf. Ser. **189**, 012003 (2009) [arXiv:0812.4703 [gr-qc]].
A. Corichi and P. Singh, A geometric perspective on singularity resolution and uniqueness in loop quantum cosmology, Phys. Rev. D **80**, 044024 (2009) [arXiv:0905.4949 [gr-qc]].
M. Bojowald, Quantum gravity effects on space-time, arXiv:1002.2618 [gr-qc].

[42] L. Modesto, Disappearance of black hole singularity in quantum gravity, Phys. Rev. D **70**, 124009 (2004) [arXiv:gr-qc/0407097].
M. Bojowald, R. Goswami, R. Maartens and P. Singh, A black hole mass threshold from non-singular quantum gravitational collapse, Phys. Rev. Lett. **95** (2005) 091302 [arXiv:gr-qc/0503041].

[43] T. Thiemann, Anomaly-free formulation of non-perturbative, four-dimensional Lorentzian quantum gravity, Phys. Lett. B **380** (1996) 257 [arXiv:gr-qc/9606088]. T. Thiemann, Quantum spin dynamics (QSD), Class. Quant. Grav. **15** (1998) 839 [arXiv:gr-qc/9606089].

[44] L. Smolin, The classical limit and the form of the Hamiltonian constraint in non-perturbative quantum general relativity, arXiv:gr-qc/9609034.

[45] T. Thiemann, "Quantum spin dynamics. VIII: The master constraint," Class. Quant. Grav. **23** (2006) 2249 [arXiv:gr-qc/0510011].

[46] C. Rovelli, "A new look at loop quantum gravity," arXiv:1004.1780 [gr-qc].

[47] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourao and T. Thiemann, "Coherent state transforms for spaces of connections," J. Funct. Anal. **135** (1996) 519 [arXiv:gr-qc/9412014].
T. Thiemann, "Gauge field theory coherent states (GCS). I: General properties," Class. Quant. Grav. **18** (2001) 2025 [arXiv:hep-th/0005233].

[48] L. Freidel and S. Speziale, "Twisted geometries: A geometric parametrisation of SU(2) phase space," arXiv:1001.2748 [gr-qc].

[49] M. Kapovich and J. J. Millson, "The symplectic geometry of polygons in Euclidean space," J. Differential Geom. **44**, 3 (1996), 479-513.

[50] A. D. Alexandrov, "Convex Polyhedra," Springer (2005).

[51] E. Bianchi, P. Doná and S. Speziale, "Polyhedra and loop quantum gravity," to appear.

[52] B. Dittrich and S. Speziale, "Area-angle variables for general relativity," New J. Phys. **10** (2008) 083006 [arXiv:0802.0864 [gr-qc]].

[53] C. Rovelli and S. Speziale, "On the geometry of loop quantum gravity on a graph," arXiv:1005.2927 [gr-qc].

[54] L. Freidel and S. Speziale, "From twistors to twisted geometries," arXiv:1006.0199 [gr-qc].

[55] L. Freidel and E. R. Livine, "U(N) Coherent States for Loop Quantum Gravity," arXiv:1005.2090 [gr-qc].

[56] MP Reisenberger, World sheet formulations of gauge theories and gravity, arXiv:gr-qc/9412035.
J Iwasaki, Geometries induced by surfaces: An algorithm on the 3-dimensional simplex lattice, in *The 7th Marcel Grossmann meeting on general relativity, Stanford 1994, Proceedings*, ed. RT Jantzen, GM Keiser, R Ruffini (River Edge, N.J., World Scientific, 1996), pp. 803-804.

[57] MP Reisenberger, C Rovelli, "Sum over surfaces" form of loop quantum gravity, *Phys. Rev.* **D56** , (1997) 3490; arXiv:gr-qc/9612035.

JC Baez, Spin foam models, *Class. Quant. Grav.* **15** (1998) 1827-1858; arXiv:gr-qc/9709052.
MP Reisenberger, C Rovelli, Spin foams as Feynman diagrams, in *Florence 2001, A relativistic spacetime odyssey* ed. I Ciufolini, D Dominici, L Lusanna (World Scientific, Singapore, 2003) pp. 431-448; arXiv:gr-qc/0002083.

JC Baez, An introduction to spin foam models of BF theory and quantum gravity, *Lect. Notes Phys.* **543** (2000) 25-94.

MP Reisenberger, C Rovelli, Spacetime as a Feynman diagram: The connection formulation, *Class. Quant. Grav.* **18** (2001) 121-140.

[58] MP Reisenberger, "A lattice worldsheet sum for 4-d Euclidean general relativity", arXiv:gr-qc/9711052.

[59] MP Reisenberger, "A lattice worldsheet sum for 4-d Euclidean general relativity", arXiv:gr-qc/9711052.

[60] JW Barrett, L Crane, "Relativistic spin networks and quantum gravity", *J. Math. Phys.* **39** (1998) 3296-3302.
JW Barrett, L Crane, "A Lorentzian signature model for quantum general relativity", *Class. Quantum Grav.* **17** (2000) 3101-3118.

[61] R DePietri, L Freidel, K Krasnov, C Rovelli "Barrett-Crane model from a Boulatov-Ooguri field theory over a homogeneous space", *Nucl. Phys.* B**574** (2000) 785-806.
A Perez, C Rovelli, "A spinfoam model without bubble divergences", *Nucl. Phys.* **B599** (2001) 255-282.
D Oriti, RM Williams, "Gluing 4-simplices: a derivation of the Barrett-Crane spinfoam model for Euclidean quantum gravity", *Phys. Rev.* **D63** (2001) 024022.

[62] D Oriti, "Spacetime geometry from algebra: spin foam models for non-perturbative quantum gravity", *Rept. Prog. Phys.* **64** (2001) 1489-1544.
A Perez, "Spin Foam Models for Quantum Gravity", *Class. Quant. Grav.* **20** (2003) R43.

[63] K Noui, A Perez, "Three dimensional loop quantum gravity: Physical scalar product and spin foam models", *Class. Quant. Grav.* **22** (2005) 1739; arXiv:gr-qc/0402110.

[64] A Ashtekar, M Campiglia and A Henderson, "Loop Quantum Cosmology and Spin Foams", *Phys. Lett.* **B681** (2009) 347-352; arXiv:0909.4221 [gr-qc].
A Ashtekar, M Campiglia and A Henderson, "Casting Loop Quantum Cosmology in the Spin Foam Paradigm", arXiv:1001.5147 [gr-qc].

[65] MV Battisti, A Marciano and Carlo Rovelli, "Triangulated Loop Quantum Cosmology: Bianchi IX and inhomogenous perturbations", arXiv:0911.2653 [gr-qc]
C Rovelli, F Vidotto, " On the spinfoam expansion in cosmology", arXiv:0911.3097 [gr-qc]
E Bianchi, C Rovelli and F Vidotto, " Towards Spinfoam Cosmology", arXiv:1003.3483 [gr-qc]

[66] J Engle, R Pereira, C Rovelli, "The loop-quantum-gravity vertex-amplitude", *Phys. Rev. Lett.* **99** (2007) 161301; arXiv:0705.2388 [gr-qc].

[67] J Engle, R Pereira, and C Rovelli, "Flipped spinfoam vertex and loop gravity", *Nucl. Phys.* **B798** (2008) 251-290; arXiv:0708.1236 [gr-qc].

[68] J Engle, E Livine, R Pereira, and C Rovelli, "LQG vertex with finite Immirzi parameter", *Nucl. Phys.* **B799** (2008) 136-149; arXiv:0711.0146[gr-qc].

[69] R Pereira, "Lorentzian LQG vertex amplitude" , *Class. Quant. Grav.* **25** (2008) 085013; arXiv:0710.5043 [gr-qc]

[70] E Livine and S Speziale, "A new spinfoam vertex for quantum gravity", *Phys. Rev.* **D76** (2007) 084028; arXiv:0705.0674 [gr-qc].

[71] L Freidel and K Krasnov, "A New Spin Foam Model for 4d Gravity", *Class. Quant. Grav.* **25**, 125018 (2008); arXiv:0708.1595 [gr-qc].

[72] S Alexandrov, "The new vertices and canonical quantization", arXiv:1004.2260[gr-qc].

[73] E Bianchi, D Regoli, C Rovelli, "Face amplitude of spinfoam quantum gravity", arXiv:1005.0764 [gr-qc].

[74] W Ruhl, *The Lorentz group and harmonic analysis* (W.A. Benjamin, Inc., New York, 1970).
IM Gel'fand, MI Graev, NYa Vilenkin, *Generalized Functions: Volume 5 Integral Geometry and Representation Theory* (Academic Press, 1966).

[75] IM Gel'fand, RA Minlos and Z Ya Shapiro, *Representations of the rotation and Lorentz groups and their applications* (Pergamon Press, 1963), pp. 187-189.

[76] T Regge, "General Relativity Without Coordinates", *Nuovo Cimento* **19** (1961) 558-571.

[77] A. Barbieri, "Quantum tetrahedra and simplicial spin networks," Nucl. Phys. B **518** (1998) 714 [arXiv:gr-qc/9707010].

[78] J. C. Baez and J. W. Barrett, "The quantum tetrahedron in 3 and 4 dimensions," Adv. Theor. Math. Phys. **3** (1999) 815 [arXiv:gr-qc/9903060].

[79] L. Freidel and S. Speziale, "Twisted geometries: A geometric parametrisation of SU(2) phase space," arXiv:1001.2748 [gr-qc].
L. Freidel and S. Speziale, "From twistors to twisted geometries," arXiv:1006.0199 [gr-qc].

[80] J. Roberts, "Classical 6j-symbols and the tetrahedron," Geom. Topol. **3** (1999), 21-66 [arXiv:math-ph/9812013].

[81] G. P. Michon, `http://www.numericana.com/data/polyhedra.htm`

[82] L. Charles, "On the quantization of polygon spaces," [arXiv:0806.1585].

[83] L. Freidel, K. Krasnov and E. R. Livine, "Holomorphic Factorization for a Quantum Tetrahedron," arXiv:0905.3627 [hep-th].

[84] M. Kapovich and J. Millson, "Quantization of bending deformations of polygons in E3 , hypergeometric integrals and the Gassner representation," Canad. Math. Bull., Vol. 44, (2001) p. 36-60

[85] H. S. M. Coxeter, *Regular Polytopes*, (3rd edition, 1973), Dover.

[86] J. B. Lasserre, "An analytical expression and an algorithm for the volume of a Convex Polyhedron in Rn", J. Optim. Theor. Appl. 39, pp. 363–377.

[87] U. Pachner, "PL homeomorphic manifolds are equivalent by elementary shellings," European Journal of Combinatorics **12**, 129 (1991).

[88] G. Immirzi, "Quantizing Regge calculus," Class. Quant. Grav. **13** (1996) 2385 [arXiv:gr-qc/9512040].
G. Immirzi, "Quantum gravity and Regge calculus," Nucl. Phys. Proc. Suppl. **57** (1997) 65 [arXiv:gr-qc/9701052].

[89] B. Dittrich and J. P. Ryan, "Phase space descriptions for simplicial 4d geometries," `arXiv:0807.2806`.

[90] V. Bonzom, "From lattice BF gauge theory to area-angle Regge calculus," *Class. Quant. Grav.* **26** (2009) 155020, `arXiv:0903.0267`.

[91] D. Oriti and T. Tlas, "Encoding simplicial quantum geometry in group field theories," `arXiv:0912.1546`.

[92] K. Giesel and T. Thiemann, "Consistency check on volume and triad operator quantisation in loop quantum gravity. I," Class. Quant. Grav. **23** (2006) 5667 [arXiv:gr-qc/0507036].

[93] T. Thiemann, "Gauge field theory coherent states (GCS). I: General properties," Class. Quant. Grav. **18** (2001) 2025 [arXiv:hep-th/0005233].
T. Thiemann and O. Winkler, "Gauge field theory coherent states (GCS) III: Ehrenfest theorems," Class. Quant. Grav. **18** (2001) 4629 [arXiv:hep-th/0005234].
T. Thiemann and O. Winkler, "Gauge field theory coherent states (GCS). II: Peakedness properties," Class. Quant. Grav. **18** (2001) 2561 [arXiv:hep-th/0005237].
H. Sahlmann, T. Thiemann and O. Winkler, "Coherent states for canonical quantum general relativity and the infinite tensor product extension," Nucl. Phys. B **606** (2001) 401 [arXiv:gr-qc/0102038].
T. Thiemann, "Complexifier coherent states for quantum general relativity," Class. Quant. Grav. **23** (2006) 2063 [arXiv:gr-qc/0206037].

[94] B. Bahr and T. Thiemann, "Gauge-invariant coherent states for Loop Quantum Gravity II: Non-abelian gauge groups," Class. Quant. Grav. **26** (2009) 045012 [arXiv:0709.4636 [gr-qc]].

[95] C. Flori and T. Thiemann, "Semiclassical analysis of the Loop Quantum Gravity volume operator: I. Flux Coherent States," arXiv:0812.1537 [gr-qc].

[96] E. Bianchi, E. Magliaro and C. Perini, "Coherent spin-networks," arXiv:0912.4054 [gr-qc].

[97] L. Freidel and S. Speziale, "Quantum twisted geometries and coherent states," to appear.

[98] B.C. Hall, "Geometric quantization and the generalized SegalBargmann transclass for Lie groups of compact type," Comm. Math. Phys. **226** (2002) 233268.

[99] K. Kowalski, J. Rembielinski and L. C. Papaloucas, "Coherent states for a quantum particle on a circle," J. Phys. A **29** (1996) 4149 [arXiv:quant-ph/9801029].

[100] E. R. Livine and S. Speziale, "A new spinfoam vertex for quantum gravity," Phys. Rev. D **76** (2007) 084028 [arXiv:0705.0674 [gr-qc]].

[101] E. R. Livine and S. Speziale, "Consistently Solving the Simplicity Constraints for Spinfoam Quantum Gravity," Europhys. Lett. **81** (2008) 50004 [arXiv:0708.1915 [gr-qc]].

[102] F. Conrady and L. Freidel, "Quantum geometry from phase space reduction," arXiv:0902.0351 [gr-qc].

[103] J. W. Barrett, R. J. Dowdall, W. J. Fairbairn, H. Gomes and F. Hellmann, "Asymptotic analysis of the EPRL four-simplex amplitude," arXiv:0902.1170 [gr-qc].

[104] J. W. Barrett, R. J. Dowdall, W. J. Fairbairn, F. Hellmann and R. Pereira, "Lorentzian spin foam amplitudes: graphical calculus and asymptotics," arXiv:0907.2440 [gr-qc].

[105] L. Freidel and E. R. Livine, "The Fine Structure of SU(2) Intertwiners from U(N) Representations," arXiv:0911.3553 [gr-qc].

[106] B. Dittrich and S. Speziale, "Area-angle variables for general relativity," New J. Phys. **10** (2008) 083006 [arXiv:0802.0864 [gr-qc]].

[107] B. Dittrich and J. P. Ryan, "Phase space descriptions for simplicial 4d geometries," arXiv:0807.2806 [gr-qc].

[108] E. Bianchi, "The length operator in Loop Quantum Gravity," Nucl. Phys. B **807** (2009) 591 [arXiv:0806.4710 [gr-qc]].

[109] H. Waelbroeck and J. A. Zapata, "A Hamiltonian formulation of topological gravity," Class. Quant. Grav. **11** (1994) 989 [arXiv:gr-qc/9311035].

[110] S. A. Major, "Shape in an Atom of Space: Exploring quantum geometry phenomenology," arXiv:1005.5460 [gr-qc].

[111] P. Gritzmann and V. Klee, On the complexity of some basic problems in computational convexity: II. Volume and mixed volumes, Polytopes: Abstract, Convex and Computational (Boston) (T. Bisztriczky, P. McMullen, R. Schneider, and A. I. Weiss, eds.), Kluwer, 1994, pp. 373-466.

[112] B. Büeler and A. Enge and K. Fukuda, "Exact volume computation for polytopes: A practical study.", Polytopes - Combinatorics and Computation, DMV-Seminars vol. 29., Birkhäuser Verlag, Basel 2000, pp. 131–154.

[113] J.J. Little, "Extended Gaussian images, mixed volumes, shape reconstruction", SCG '85: Proceedings of the first annual symposium on Computational geometry, pp. 15–23

[114] Godsil, C. and Royle, G. (2001). Algebraic Graph Theory. Springer Verlag.

[115] L. Freidel and E. R. Livine, "The Fine Structure of SU(2) Intertwiners from U(N) Representations," arXiv:0911.3553 [gr-qc].

[116] L. Freidel and E. R. Livine, "U(N) Coherent States for Loop Quantum Gravity," arXiv:1005.2090 [gr-qc].

[117] E. F. Borja, J. Diaz-Polo, I. Garay and E. R. Livine, "Dynamics for a 2-vertex Quantum Gravity Model," arXiv:1006.2451 [gr-qc].

[118] L. G. Yaffe, "Large N Limits As Classical Mechanics," Rev. Mod. Phys. **54** (1982) 407.

[119] Y. Ma, C. Soo and J. Yang, "New length operator for loop quantum gravity," Phys. Rev. D **81** (2010) 124026 [arXiv:1004.1063 [gr-qc]].

[120] T. Thiemann, "Closed formula for the matrix elements of the volume operator in canonical quantum gravity," J. Math. Phys. **39** (1998) 3347 [arXiv:gr-qc/9606091].
R. De Pietri, "Spin networks and recoupling in loop quantum gravity," Nucl. Phys. Proc. Suppl. **57** (1997) 251 [arXiv:gr-qc/9701041].

[121] J. Brunnemann and D. Rideout, "Properties of the Volume Operator in Loop Quantum Gravity I: Results," Class. Quant. Grav. **25** (2008) 065001 [arXiv:0706.0469 [gr-qc]].

[122] R. J. Glauber, "Coherent and incoherent states of the radiation field," Phys. Rev. **131** (1963) 2766.

[123] E. C. G. Sudarshan, "Equivalence of semiclassical and quantum mechanical descriptions of, statistical light beams," Phys. Rev. Lett. **10** (1963) 277.

[124] A. Ashtekar and J. Lewandowski, "Quantum theory of geometry. II: Volume operators," Adv. Theor. Math. Phys. **1**, 388 (1998) [arXiv:gr-qc/9711031].

[125] W. Kaminski, M. Kisielowski and J. Lewandowski, "Spin-Foams for All Loop Quantum Gravity," Class. Quant. Grav. **27** (2010) 095006 [arXiv:0909.0939 [gr-qc]].

[126] E. Bianchi, E. Magliaro and C. Perini, "Spinfoams in the holomorphic representation," arXiv:1004.4550 [gr-qc].

[127] E. Bianchi, C. Rovelli and F. Vidotto, "Towards Spinfoam Cosmology," arXiv:1003.3483 [gr-qc].

[128] E. Bianchi, E. Magliaro and C. Perini, "LQG propagator from the new spin foams," Nucl. Phys. B **822** (2009) 245 [arXiv:0905.4082 [gr-qc]].

[129] A. M. Perelomov, *Generalized Coherent States and Their Applications* (Springer-Verlag, 1986).

[130] F. A. Berezin, "Non-wiener continual integrals", Teor. Mat. Fiz. **6** (1971) 194.

[131] F. A. Berezin, "Feynman Path Integrals In A Phase Space," Sov. Phys. Usp. **23** (1981) 763 [Usp. Fiz. Nauk **132** (1980) 497].

[132] J. W. Barrett and L. Crane, "Relativistic spin networks and quantum gravity," J. Math. Phys. **39** (1998) 3296 [arXiv:gr-qc/9709028].

[133] F. Conrady and L. Freidel, "On the semiclassical limit of 4d spin foam models," Phys. Rev. D **78** (2008) 104023 [arXiv:0809.2280 [gr-qc]].